Convergence analysis of monotone schemes for second-order non-linear parabolic PDEs and their applications in sublinear expectation

by

Shuo Huang

Thesis

Submitted to the University of Warwick

for admission to the degree of

Doctor of Philosophy

Department of Statistics

Oct 2020

Contents

List of Figures

Acknowledgments

I would like to thank my supervisor, Dr. Gechun Liang, for providing consistent advice, guidance and support throughout my period of PhD study. He originally enrolled me as his PhD student when he was at King's College London in 2016, and later transferred me together with him to the University of Warwick in 2017. Dr. Liang organized a weekly reading group during my first two years of study, which led me to exciting areas of research and giving me huge help and inspiration. He has always been a great mentor, and given me dedicated support not just in my research life but in my personal life and career development.

I'm also grateful to Dr. Martin Herdegen, who organized successful seminars and the reading group within Stochastic Finance group in which a great number of interesting topics were introduced. My huge thanks go to other students who presented in the reading group and shared their understanding. Those names include, but are not limited to, Haodong Sun, Matthew Zeng, Dominic Norgilas, Jonathan Muscat, Joseph Jerome and Kevin Engelbrecht.

I would also like to express my sincere gratitude to my beloved partner Yao, for the constant love and support, and always encouraging me when I was in low spirit. I also want to thank my parents for their love and kind financial support, especially during my first year at King's.

Finally, I acknowledge funding from the Department of Statistics since my second year.

Declarations

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree. The work presented was carried out by the myself except where stated otherwise by references.

- Chapter 3 is a joint work with Gechun Liang and Thaleia Zariphopoulou and has recently been published at:

  S. Huang, G. Liang, and T. Zariphopoulou. An approximation scheme for semilinear parabolic PDEs with convex and coercive Hamiltonians. *SIAM Journal on Control and Optimization*, 2020, Vol. 58, No. 1 : pp. 165-191.

- Chapter 5 and 6 are based on the following manuscript which was submitted for publication:

  S. Huang, G. Liang. A monotone scheme for G-equations with application to the explicit convergence rate of robust central limit theorem. *arXiv preprint arXiv:1904.07184,* 2020.

Abstract

Second-order non-linear parabolic partial differential equations have been a central research area for decades due to their various applications to physics, engineering, finance and others. To solve such equations, numerical approximations are usually designed to converge to a weak version of solutions called viscosity solutions. This thesis proposes monotone approximation schemes for two specific types of non-linear parabolic equations arising in applied mathematics, and establishes the convergence and convergence rate to viscosity solutions. The proposed schemes involve only semi-discretization in time and the rates of convergence are represented in the form of some exponent of the time step scaled by a constant.

We first propose a monotone scheme for a class of semi-linear parabolic equations that are convex and coercive in their gradients arising from utility indifference pricing in mathematical finance. The proposed scheme is based on a splitting method and its convergence rate is determined by combining Krylov's shaking coefficients technique and the Barles-Jakobsen optimal switching approximation. An extension to variational inequalities is also studied using an obstacle switching system. We then build a piece-wise constant monotone scheme for a class of fully non-linear equations called G-equations arising from Knightian uncertainty in statistics, and determine its convergence rate with an explicit error bound. We present three applications under a sublinear expectation framework: a convergence rate for Peng's robust central limit theorem with an explicit bound of Berry-Esseen type, a monoton scheme for the Black-Scholes-Barenblatt (BSB) equation that is a natural generalization of Cox-Ross-Rubinstein (CIR) binomial tree approximation to the case with model ambiguity, and an optimal switching approximation to G-normal distribution.

Chapter 1

Introduction

1.1 A brief history of numerical approximation to viscosity solutions and sublinear expectation

The notion of viscosity solution was originally introduced by Crandall and Lions [19] for first-order Hamilton-Jacobi equations. Since then, numerous existence and uniqueness results of viscosity solutions of Hamilton-Jacobi equations have been obtained (see [17, 20, 36] and more references therein). The notion of viscosity solution was later extended by Lions [46] to second-order Hamilton-Jacobi-Bellman (HJB) equations with connection to optimal control of diffusion processes, and by Ishii [38] to second-order Hamilton-Jacobi-Bellman-Isaacs (HJBI) equations which connects to differential games. For a detailed introduction of viscosity solutions and the existence and uniqueness results of second-order nonlinear PDEs, we refer to the User's guide [18].

Numerical approximation schemes for viscosity solutions were first studied by Barles and Souganidis [5], who showed that any monotone, stable and consistent approximation scheme converges to the correct solution, provided that there exists a comparison principle for the limiting equation. The corresponding convergence rate had been an open problem for a long time until late 1990s when Krylov [43, 44] introduced the shaking coefficients technique to construct a sequence of smooth subsolutions/supersolutions. This technique was further developed by Barles and Jacobsen in a sequence of papers (see [2] and [39] and more references therein), and has recently been applied to solve various problems (see, among others, [6] [9] [28] and [33]).

Krylov's technique depends crucially on the convexity/concavity of the underlying equation with respect to its terms. As a result, unless the approximate solution has enough regularity (so one can interchange the roles of the approximation scheme and the original equation), the shaking coefficients technique only gives either an upper or a lower bound for the approximation error, but not both. A further breakthrough was made by Barles and Jacobsen in [3] and [4], who combined the ideas of optimal switching approximation of Hamilton-Jacobi-Bellman (HJB) equations (initially proposed by Evans and Friedman [27]) with the shaking coefficients technique. They obtained both upper and lower bounds of the error estimate, but with a lower convergence rate due to the introduction of another approximation layer.

On the other hand, the framework of sublinear expectation was introduced by Peng in a series of paper [51–54] where he generalized the classical probability theory in the presence of Knightian uncertainty, for problems such as model uncertainty in statistics, measure of risk and superhedging in finance. The basic idea is that the expectation of random variables is evaluated by a supremum over a range of classical linear expectations under different probability measure, and thus it is sublinear. Peng introduced G-normal distribution and characterized G-normal random variables by a class of fully nonlinear PDEs called G-equations, which are inherently connected to HJB equations. A fundamental result about G-normal distribution is the central limit theorem under sublinear expectation framework and was first proved in [52]. The corresponding convergence rate was subsequently obtained by Song [60] and Fang et al [30] using Stein's method and more recently by Krylov [45] using stochastic control method under different model assumptions.

1.2 Outline

The rest of this thesis is organized as follows. We provide some background material of the thesis in Chapter 2. The basic theory of viscosity solutions for second-order partial differential equations and the Barles-Souganidis convergence framework are first introduced. We then present a standard convergence analysis in the case of linear parabolic equations and apply it to the convergence rate of classical central limit theorem.

The first main part of the thesis is given in Chapter 3-4. In Chapter 3 we propose an approximation scheme for a class of semilinear parabolic equations with convex and coercive Hamiltonians. We prove its convergence and determine the convergence rate, combining Krylov's shaking coefficients technique and Barles-Jakobsen's optimal switching approximation. Chapter 4 then extends the work in Chapter 3 by considering a variational inequality version of the PDE in Chapter 3. We propose an adapted approximation scheme based on the scheme used in Chapter 3 and apply a similar procedure to establish the convergence rate. However, a main difficulty arise when applying optimal switching approximation method. We introduce an obstacle switching system to tackle this difficulty.

The second part of the thesis is devoted to the fully nonlinear G-equations under the sublinear expectation framework. We give some theoretical results in Chapter 5, by proposing a piece-wise constant approximation scheme for G-equations and establishing its convergence rate with an explicit approximation error bound. Chapter 6 then introduces some applications of the theoretical convergence results under sublinear expectation framework. In particular, we obtain an explicit convergence rate of Berry-Esseen type for the law of large numbers and central limit theorem under a sublinear expectation.

Finally, Chapter 7 summarizes the main contribution of the thesis and discusses some potential directions of future work.

Chapter 2

Background

This chapter is devoted to some background material of this thesis. We first establish general notation and then introduce the basic theory of viscosity solutions for second-order partial differential equations along with some useful properties. We next introduce the famous Barles-Souganidis framework which states that any monotone, stable and consistent approximation scheme converges to the unique continuous viscosity solution, given that a strong comparison principle holds.

However, this framework achieves convergence only, and is unable to estimate the convergence rate. To give a simple example and an idea of how to estimate scheme convergence rates, we present later a standard convergence analysis in the case of linear parabolic equations with finite difference being the approximation scheme. This, on the other hand, shows some difficulties we would face when considering nonlinear equations, which are the main topic of the following chapters of this thesis. Finally, we give a direct application to the convergence rate of classical *Central Limit Theorem* (CLT).

2.1 General Notation

In this section, we introduce general notation that will be used throughout this thesis.

Euclidean norm. Let $d, d' > 0$ be two integers. For any $x \in \mathbb{R}^d$, we denote $|x|$ by its Euclidean norm:

$$|x| := \sqrt{x^T x}.$$

We regard $d \times d'$ matrices as $\mathbb{R}^{d \times d'}$ vectors. Note that for such matrices M, $|M|^2 = \mathrm{tr}(M^T M)$.

Function Spaces. Let $d, m > 0, k \geq 0$ be three integers and $0 < \delta \leq 1$. For a function $f : \Omega \subset \mathbb{R}^d \to \mathbb{R}^m$, we define its (semi)norms

$$|f|_0 := \sup_{x \in \Omega} |f(x)|, \quad [f]_{\mathcal{C}^\delta} := \sup_{x \neq x' \in \Omega} \frac{|f(x) - f(x')|}{|x - x'|^\delta},$$

and let

$$|f|_\delta := |f|_0 + [f]_{\mathcal{C}^\delta}.$$

Furthermore, if f is differentiable up to order k, we define

$$|f|_{\mathcal{C}^{k+\delta}} := \max_{|\beta|\le k} |D^\beta f|_0 + \max_{|\beta|=k} [D^\beta f]_{\mathcal{C}^\delta},$$

where β is a d-dimensional multi-index. Note that $|f|_\delta = |f|_{\mathcal{C}^{0+\delta}}$.

We then denote

- $\mathcal{C}(\Omega)$: the space of continuous real-valued functions on Ω.
- $\mathcal{C}_b(\Omega)$: the space of bounded continuous real-valued functions on Ω with finite norm $|f|_0$.
- $\mathcal{C}_b^\delta(\Omega)$: the space of bounded continuous real-valued functions on Ω with finite norm $|f|_\delta$.
- $\mathcal{C}_b^{k+\delta}(\Omega)$: the space of bounded continuous real-valued functions on Ω with finite norm $|f|_{\mathcal{C}^{k+\delta}}$.
- $\mathcal{C}_b^\infty(\Omega)$: the space of bounded smooth real-valued functions on Ω with bounded derivatives of any order.

For parabolic problems, we need to consider functions of both time and space. Let $j \in \mathbb{N}$ and $0 < \gamma \le 1$, and let $\mathcal{T}$ be some time interval in $\mathbb{R}$. For a function $f : \mathcal{T} \times \Omega \to \mathbb{R}^m$, we define its (semi)norms

$$|f|_0 := \sup_{(t,x)\in\mathcal{T}\times\Omega} |f(t,x)|, \quad [f]_{\mathcal{C}^{\gamma,\delta}} := \sup_{(t,x)\neq(t',x')\in\mathcal{T}\times\Omega} \frac{|f(t,x) - f(t',x')|}{|t-t'|^\gamma + |x-x'|^\delta},$$

and let

$$|f|_\delta := |f|_0 + [f]_{\mathcal{C}^{\delta/2,\delta}}.$$

Furthermore, if f is differentiable with respect to t up to order j and with respect to x up to order k, we define

$$|f|_{\mathcal{C}^{j+\gamma,k+\delta}} := \max_{\alpha\le j, |\beta|\le k} |\partial_t^\alpha D_x^\beta f|_0 + \max_{|\beta|=k} [\partial_t^j D_x^\beta f]_{\mathcal{C}^{\gamma,\delta}},$$

and let $\mathcal{C}_b^{j+\gamma,k+\delta}(\mathcal{T}\times\Omega)$ denote the space of bounded continuous real-valued functions on $\mathcal{T}\times\Omega$ with finite norm $|f|_{\mathcal{C}^{j+\gamma,k+\delta}}$.

L^p Spaces. Let d be a positive integer and $1 \le p < \infty$. For a function $f : \mathbb{R}^d \to \mathbb{R}$, we define its L^p norm

$$\|f\|_{L^p} := \left(\int_{\mathbb{R}^d} |f(x)|^p dx \right)^{1/p},$$

and we simply use $\|f\|_p$ to denote $\|f\|_{L^p}$.

Deterministic/Stochastic Processes Spaces. Let $1 \le p < \infty$ and $0 \le t < T \le \infty$, we denote

- $\mathbb{L}^p[t,T]$: the space of deterministic process q_t such that $\int_t^T |q_s|^p ds < \infty$.
- $\mathbb{H}^p[t,T]$: the space of stocastic process q_t such that $\mathbb{E}\left[\int_t^T |q_s|^p ds\right] < \infty$.

2.2 Viscosity Solutions

We consider d-dimensional second order degenerate partial differential equations of the form

$$F(x, u(x), Du(x), D^2u(x)) = 0 \tag{2.1}$$

for $x \in \mathcal{O} \subset \mathbb{R}^d$, where $\mathcal{O}$ is an open set and $F : \mathcal{O} \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}_d \to \mathbb{R}$ is assumed to be continuous and $\mathcal{S}_d$ is the set of symmetric $d \times d$ matrices.

A fundamental *degenerate ellipticity* condition is needed throughout the theory of viscosity solutions:

Assumption 2.2.1 *For any $(x, r, p) \in \mathcal{O} \times \mathbb{R} \times \mathbb{R}^d$ and $X, Y \in \mathcal{S}_d$,*

$$F(x, r, p, X) \leq F(x, r, p, Y) \quad \textit{whenever} \quad X \geq Y. \tag{2.2}$$

To give the definition of viscosity solutions, we need the following notations: for a locally bounded function $u : \mathcal{O} \to \mathbb{R}$,

$$u^*(x) = \limsup_{x' \to x} u(x'), \quad u_*(x) = \liminf_{x' \to x} u(x').$$

They are called upper- and lower-semicontinuous envelope of u respectively. Recall that u^* is the smallest upper-semicontinuous function greater than u, and u_* is the largest lower-semicontinuous function less than u.

Definition 2.2.2 *Let $u : \mathcal{O} \to \mathbb{R}$ be locally bounded, then:*
(i) u is a viscosity subsolution of (2.1) if

$$F(x_0, u^*(x_0), D\varphi(x_0), D^2\varphi(x_0)) \leq 0$$

for any pair $(x_0, \varphi) \in \mathcal{O} \times \mathcal{C}^2(\mathcal{O})$ such that x_0 is a (local) maximizer of $(u^* - \varphi)$ on $\mathcal{O}$.
(ii) u is a viscosity supersolution of (2.1) if

$$F(x_0, u_*(x_0), D\varphi(x_0), D^2\varphi(x_0)) \geq 0$$

for any pair $(x_0, \varphi) \in \mathcal{O} \times \mathcal{C}^2(\mathcal{O})$ such that x_0 is a (local) minimizer of $(u_* - \varphi)$ on $\mathcal{O}$.
(iii) u is a viscosity solution of (2.1) if it is both viscosity subsolution and viscosity supersolution of (2.1).

To simplify the definition, we introduce the following notations which are commonly called *semijets* in literature:

Definition 2.2.3 *For any $\hat{x} \in \mathcal{O}$, the superjet $J^{2,+}_{\mathcal{O}} u(\hat{x})$ of a USC function u on $\mathcal{O}$ is defined by*

$$J^{2,+}_{\mathcal{O}} u(\hat{x}) := \left\{ (D\varphi(\hat{x}), D^2\varphi(\hat{x})) : \varphi \in \mathcal{C}^2(\mathcal{O}) \textit{ and } \hat{x} \textit{ is a local maximizer of } u - \varphi \right\}.$$

Similarly, the subjet $J^{2,-}_{\mathcal{O}} u(\hat{x})$ of a LSC function u on $\mathcal{O}$ is defined by

$$J^{2,-}_{\mathcal{O}} u(\hat{x}) := \left\{ (D\varphi(\hat{x}), D^2\varphi(\hat{x})) : \varphi \in \mathcal{C}^2(\mathcal{O}) \textit{ and } \hat{x} \textit{ is a local minimizer of } u - \varphi \right\}.$$

Further, we enlarge the semijets to their closures:

$$\bar{J}^{2,\pm}_{\mathcal{O}}u(\hat{x}) := \Big\{(p,X) : (x_n, u(x_n), p_n, X_n) \to (\hat{x}, u(\hat{x}), p, X) \text{ for some sequence } (x_n, p_n, X_n)_n \text{ such that } (p_n, X_n) \in J^{2,\pm}_{\mathcal{O}}u(x_n)\Big\}$$

Remark 2.2.4 *By Definition 2.2.2 and 2.2.3, we have u is a viscosity subsolution of (2.1) if for any $x \in \mathcal{O}$ and $(p,X) \in J^{2,+}_{\mathcal{O}}u^*(x)$,*

$$F(x, u^*(x), p, X) \leq 0.$$

This, due to the continuity of F, remains true if $(p,X) \in \bar{J}^{2,+}_{\mathcal{O}}u^*(x)$. Similar argument holds also for supersolutions and solutions.

Remark 2.2.5 *For parabolic problems, we need to consider any point in $\mathbb{R}^d$ $(d \geq 2)$ as a time variable t in $\mathbb{R}$ and a space variable x in $\mathbb{R}^{d-1}$. Further, the open set $\mathcal{O}$ is usually of the form $(0,T) \times \mathcal{O}_{d-1}$ where $T > 0$ and $\mathcal{O}_{d-1} \in \mathbb{R}^{d-1}$ is an open set, and equation (2.1) is usually written as*

$$\partial_t u + F(t, x, u(t,x), D_x u(t,x), D^2_x u(t,x)) = 0 \tag{2.3}$$

A parabolic version of semijets is denoted as $\mathcal{P}^{2,\pm}_{\mathcal{O}}$. Similar to Remark 2.2.4, we have u is a viscosity subsolution of (2.3) if for any $(t,x) \in \mathcal{O}$ and $(a,p,X) \in \bar{\mathcal{P}}^{2,+}_{\mathcal{O}}u^*(t,x)$,

$$a + F(t, x, u^*(t,x), p, X) \leq 0.$$

Similar argument holds also for supersolutions and solutions.

We next give the *stability result* and *comparison principle* of viscosity solutions. The former shows that viscosity solutions are stable under passage to limits and the latter compares viscosity sub- and supersolutions and implies directly the uniqueness of continuous viscosity solutions. Both properties are used frequently throughout this thesis. It is worth noting that the existence of viscosity solution is essentially given by *Perron's Method*, which states that a viscosity solution exists if the comparison principle holds and there is a subsolution and a supersolution that satisfy the boundary conditions. This is not the concern of this thesis and we refer to [18, 37, 38] for more details.

Proposition 2.2.6 (Stability, Theorem 6.8 in [64]) *Let u_ε be a viscosity subsolution of*

$$F_\varepsilon(x, u_\varepsilon(x), Du_\varepsilon(x), D^2 u_\varepsilon(x)) = 0 \text{ in } \mathcal{O},$$

where $(F_\varepsilon)_{\varepsilon>0}$ is a class of continuous functions satisfying the degenerate ellipticity condition. Define

$$\bar{u}(x) := \limsup_{(\varepsilon,x')\to(0,x)} u_\varepsilon(x') \quad \text{and} \quad \underline{F}(z) := \liminf_{(\varepsilon,z')\to(0,z)} F_\varepsilon(z').$$

Then, $\bar{u}$ is a upper-semicontinuous viscosity subsolution of

$$\underline{F}(x, \bar{u}(x), D\bar{u}(x), D^2\bar{u}(x)) = 0 \text{ in } \mathcal{O}.$$

A similar statement holds for supersolutions.

Unlike stability result, the comparison principle of viscosity solutions does not hold in general by merely degenerate ellipticity condition. We then give the definition for the comparison principle, followed by some general sufficient conditions for the comparison principle to hold.

Definition 2.2.7 *Comparison principle for equation (2.1) is the following statement: Suppose u and v are viscosity subsolution and supersolution of (2.1) respectively and $u^* \le v_*$ on $\partial\mathcal{O}$, then $u^* \le v_*$ in $\bar{\mathcal{O}}$.*

Proposition 2.2.8 (Theorem 6.17 and 6.21 in [64]) *Suppose $\mathcal{O}$ is bounded, then the comparison principle holds for (2.1) when F satisfies the following two conditions:*

(i) There exists a constant $\gamma > 0$ such that

$$F(x,r,p,X) - F(x,s,p,X) \ge \gamma(r-s) \quad \text{for } r \ge s, (x,p,X) \in \mathcal{O}\times\mathbb{R}^d\times\mathcal{S}_d. \tag{2.4}$$

(ii) There exists a function $\omega : \mathbb{R}^+ \to \mathbb{R}^+$ with $\lim_{x\to 0^+}\omega(x) = 0$ such that

$$F(y,r,\alpha(x-y),Y) - F(x,r,\alpha(x-y),X) \le \omega(\alpha|x-y|^2 + |x-y|) \tag{2.5}$$

for $x, y \in \mathcal{O}, r \in \mathbb{R}$ and $X, Y \in \mathcal{S}_d$ satisfying

$$-3\alpha \begin{pmatrix} I_d & 0 \\ 0 & I_d \end{pmatrix} \le \begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \le 3\alpha \begin{pmatrix} I_d & -I_d \\ -I_d & I_d \end{pmatrix}. \tag{2.6}$$

Moreover, if $\mathcal{O}$ is unbounded, then comparison principle holds if $|u^*(x)| + |v_*(x)| = o(|x|^2)$ as $|x| \to \infty$, and F is uniformly continuous satisfying the above two conditions with (2.6) changed to

$$-4\alpha \begin{pmatrix} I_d & 0 \\ 0 & I_d \end{pmatrix} \le \begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \le 4\alpha \begin{pmatrix} I_d & -I_d \\ -I_d & I_d \end{pmatrix}.$$

Remark 2.2.9 *Clearly, suppose comparison principle holds and u, v are two continuous viscosity solutions such that $u = v$ on $\partial\mathcal{O}$, then $u = v$ in $\bar{\mathcal{O}}$. In other words, continuous viscosity solution is unique given that comparison principle holds.*

The above condition (ii) implies the degenerate ellipticity condition (2.2). For more details, please refer to [18, 64] and the references therein.

Remark 2.2.10 *For parabolic problems as in Remark 2.2.5, the boundary $\partial\mathcal{O}$ takes the form of $(\{0\}\times\bar{\mathcal{O}}_{d-1}) \cup ((0,T)\times\partial\mathcal{O}_{d-1})$, and the comparison principle holds for (2.3) when F satisfies the same two conditions (2.4) and (2.5) for each fixed $t \in (0,T)$, with the same constant γ and function w.*

Moreover, in this thesis $\mathcal{O}_{d-1}$ is always equal to $\mathbb{R}^{d-1}$ and thus $\partial\mathcal{O}$ reduces to $\{0\}\times\mathbb{R}^{d-1}$, which corresponds to initial conditions.

We finish this section by introducing the Crandall-Ishii's Lemma which is used to prove Proposition 2.2.8, and its parabolic version which will be used several times later in this thesis. We refer to Crandall etc. [18] Theorem 3.2 and Theorem 8.3 (the parabolic version) for their technical proof.

Lemma 2.2.11 (Theorem 3.2 in [18]) *Let $u_1, u_2 \in USC(\mathcal{O})$, $\varphi \in \mathcal{C}^2(\mathcal{O}^2)$ and set $w(x) = u_1(x_1) + u_2(x_2)$ for $x = (x_1, x_2) \in \mathcal{O}^2$. Suppose $\hat{x} = (\hat{x}_1, \hat{x}_2) \in$*

$\mathcal{O}^2$ *is a local maximizer of* $w-\varphi$*, then for each* $\varepsilon>0$ *there exist* $X_1, X_2 \in \mathcal{S}_d$ *such that*

$$(D_{x_i}\varphi(\hat{x}), X_i) \in \bar{J}^{2,+}_{\mathcal{O}} u_i(\hat{x}_i) \quad \textit{for } i=1,2,$$

and

$$-\left(\varepsilon^{-1}+||A||\right) I_{2d} \le \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \le A+\varepsilon A^2$$

where $A = D^2\varphi(\hat{x}) \in \mathcal{S}_{2d}$*, and* $||A|| = \sup\{|\langle A\xi,\xi\rangle| : |\xi| \le 1\}$.

Lemma 2.2.12 (The parabolic version, Theorem 8.3 in [18]) *Let* $u_1, u_2 \in USC((0,T)\times\mathcal{O})$*,* $\varphi \in \mathcal{C}^{1,2}((0,T)\times\mathcal{O}^2)$ *and set* $w(t,x) = u_1(t,x_1)+u_2(t,x_2)$ *for* $t\in(0,T)$ *and* $x=(x_1,x_2)\in\mathcal{O}^2$*. Suppose that*

(i) $(\hat{t},\hat{x}) = (\hat{t},\hat{x}_1,\hat{x}_2) \in (0,T)\times\mathcal{O}^2$ *is a maximizer of* $w-\varphi$*.*

(ii) there is an $r>0$ *such that whenever* $(b_i,q_i,X_i) \in \mathcal{P}^{2,+}_{\mathcal{O}} u_i(t,x_i)$*,* $|t-\hat{t}|+|x_i-\hat{x}_i| \le r$ *and* $|u_i(t,x_i)|+|q_i|+||X_i|| \le M$ *hold for* $i=1,2$ *and some* $M>0$*, there exists a constant* C *such that*

$$b_i \le C \quad \textit{for } i=1,2.$$

Then, for each $\varepsilon>0$ *there exist* $b_1,b_2\in\mathbb{R}$ *and* $X_1,X_2\in\mathcal{S}_d$ *such that*

$$(b_i, D_{x_i}\varphi(\hat{t},\hat{x}), X_i) \in \bar{\mathcal{P}}^{2,+}_{\mathcal{O}} u_i(\hat{t},\hat{x}_i) \quad \textit{for } i=1,2,$$

$$b_1+b_2 = \varphi_t(\hat{t},\hat{x}),$$

and

$$-\left(\varepsilon^{-1}+||A||\right) I_{2d} \le \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \le A+\varepsilon A^2$$

where $A = D^2_x\varphi(\hat{t},\hat{x}) \in \mathcal{S}_{2d}$*, and* $||A|| = \sup\{|\langle A\xi,\xi\rangle| : |\xi| \le 1\}$.

2.3 Barles-Souganidis convergence framework

Barles and Souganidis [5] considers the equation (2.1) with a possible boundary condition $u=g$ on $\partial\mathcal{O}$, and write them together as

$$G(x,u,Du,D^2u) = 0 \quad \text{in } \bar{\mathcal{O}}, \tag{2.7}$$

with

$$G(x,u,Du,D^2u) = \begin{cases} F(x,u,Du,D^2u) & \text{in } \mathcal{O} \\ u-g & \text{on } \partial\mathcal{O} \end{cases}$$

A locally bounded function $u:\bar{\mathcal{O}}\to\mathbb{R}$ is called a viscosity subsolution (resp. supersolution) of (2.7) if u is a viscosity subsolution (resp. supersolution) of (2.1) in $\mathcal{O}$ and satisfies the following boundary conditions in the viscosity sense:

$$G_*(x,u,Du,D^2u) = \min\left(F_*(x,u,Du,D^2u), u-g\right) \le 0 \quad \text{on } \partial\mathcal{O} \tag{2.8}$$

(resp.

$$G^*(x,u,Du,D^2u) = \max\left(F^*(x,u,Du,D^2u), u-g\right) \ge 0 \quad \text{on } \partial\mathcal{O}). \tag{2.9}$$

In turn, a stronger notion of comparison principle is needed:

Definition 2.3.1 *We say that (2.7) satisfies the strong comparison principle for bounded solutions if for any bounded function u, v such that u is upper-semicontinuous subsolution and v is lower-semicontinuous supersolution of (2.7) respectively, we have $u \leq v$ in $\bar{\mathcal{O}}$.*

We then consider an approximation scheme for (2.7) of the form

$$S(\Delta, x, u^{\Delta}(x), u^{\Delta}(\cdot)) = 0 \ \text{ in } \bar{\mathcal{O}} \tag{2.10}$$

where $S : \mathbb{R}^+ \times \bar{\mathcal{O}} \times \mathbb{R} \times \mathcal{C}_b(\bar{\mathcal{O}})$ and the solution u^{Δ} of (2.10) is the approximation of the viscosity solution u of (2.7). Three crucial conditions satisfied by the approximation scheme in Barles-Souganidis convergence framework are:

Monotonicity. For any $(\Delta, x, r) \in \mathbb{R}^+ \times \bar{\mathcal{O}} \times \mathbb{R}$ and $u, v \in \mathcal{C}_b(\bar{\mathcal{O}})$,

$$S(\Delta, x, r, u) \leq S(\Delta, x, r, v) \ \text{ whenever } u \geq v.$$

Stability. For any $\Delta > 0$, the scheme (2.10) admits a solution $u^{\Delta} \in \mathcal{C}_b(\bar{\mathcal{O}})$ with the bound independent of Δ.

Consistency. For any $(x, \varphi) \in \bar{\mathcal{O}} \times \mathcal{C}_b^{\infty}(\bar{\mathcal{O}})$,

$$\limsup_{(\Delta,y,c)\to(0,x,0)} S(\Delta, y, \varphi(y) + c, \varphi + c) \leq G^*(x, \varphi(x), D\varphi(x), D^2\varphi(x)), \tag{2.11}$$

and

$$\liminf_{(\Delta,y,c)\to(0,x,0)} S(\Delta, y, \varphi(y) + c, \varphi + c) \geq G_*(x, \varphi(x), D\varphi(x), D^2\varphi(x)). \tag{2.12}$$

Theorem 2.3.2 (Theorem 2.1 in [5]) *Suppose that (2.7) satisfies the strong comparison principle for bounded solutions and the scheme (2.10) satisfies the monotonicity, stability and consistency conditions, then the scheme solution u^{Δ} converges locally uniformly to the unique continuous viscosity solution u of (2.7), as $\Delta \to 0$.*

Remark 2.3.3 *Barles and Souganidis introduces boundary conditions (2.8)(2.9) and the strong comparison principle to avoid the need to analyse the behavior of the approximation scheme when close to the boundary. However, it is sufficient to consider the usual boundary condition $u = g$ on $\partial\mathcal{O}$ and the standard comparison principle when a limit condition*

$$\lim_{(\Delta,y)\to(0,x)} u^{\Delta}(y) = g(x) \quad \textit{for } x \in \partial\mathcal{O}$$

is satisfied. An example of this case is given in Section 5.3.2.

2.4 Convergence analysis of linear parabolic PDEs

In this section, we apply the Barles-Souganidis framework to a simple linear second-order parabolic PDE and prove a finite difference scheme converges. Further, we show the classical idea and procedure of applying Krylov's *shaking the coefficients* technique to obtain a rate of convergence. To this end, we consider the following linear parabolic PDE

$$\partial_t u - \frac{1}{2} a(t, x) \partial_{xx} u = 0 \ \text{ in } (0, T] \times \mathbb{R}, \tag{2.13}$$

with initial condition

$$u(0,x) = g(x), \tag{2.14}$$

where $T > 0$ is a constant, a and g are continuous bounded real-valued functions defined on $(0,T] \times \mathbb{R}$ and $\mathbb{R}$ respectively. Further, for (2.13) to satisfy the degenerate ellipticity condition (2.2), we assume $a \geq 0$ in $(0,T] \times \mathbb{R}$. Finally, we denote $\bar{\sigma} := \sqrt{|a|_0}$ and assume $\bar{\sigma} > 0$ to avoid trivial case where $a(t,x)$ is constantly 0. The linear equation (2.13)-(2.14) is a simplified version of linear equations considered in Section 5 in [39], and a special case of Hamilton-Jacobi-Bellman (HJB) equation considered in [4]. They applied also the *shaking the coefficients* technique, and the rate of convergence result in this section can be implied by their results.

Remark 2.4.1 *The equation (2.13)-(2.14) satisfies the strong comparison principle for bounded solutions. See [5, 38] and the references therein for details.*

An explicit finite difference scheme can be written as

$$S(\Delta, t, x, u^{\Delta}(t,x), u^{\Delta}|_{t-\Delta}) = 0 \text{ in } [0,T] \times \mathbb{R} \tag{2.15}$$

where $S : (0,T) \times [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathcal{C}_b(\mathbb{R}) \to \mathbb{R}$ is defined as

$$S(\Delta, t, x, r, v) := \begin{cases} \frac{r - v(x)}{\Delta} - \frac{1}{2} a(t,x) \left[\frac{v(x - \bar{\sigma}\sqrt{\Delta}) + v(x + \bar{\sigma}\sqrt{\Delta}) - 2v(x)}{\bar{\sigma}^2 \Delta} \right] & \text{if } t \geq \Delta \\ r - g(x) & \text{if } t < \Delta \end{cases} \tag{2.16}$$

This scheme explicitly provides an iterative algorithm to compute u^{Δ}:

$$\begin{aligned} u^{\Delta}(t,x) = {} & \frac{a(t,x)}{2\bar{\sigma}^2} \left(u^{\Delta}(t - \Delta, x - \bar{\sigma}\sqrt{\Delta}) + u^{\Delta}(t - \Delta, x + \bar{\sigma}\sqrt{\Delta}) \right) \\ & + \left(1 - \frac{a(t,x)}{\bar{\sigma}^2} \right) u^{\Delta}(t - \Delta, x) \end{aligned} \tag{2.17}$$

for $t \geq \Delta$ with $u^{\Delta} = g$ for $t < \Delta$. A similar explicit finite difference scheme was used in Section 5 in [39], and they obtained the same convergence result as we will show in Theorem 2.4.13. Another type of scheme which is common in the literature is called control scheme (see [2, 13, 39]), or more generally, semi-Lagrangian scheme (see [21, 29]). Note that, when $a(t,x)$ is constant, the above finite difference scheme becomes a special case of control schemes.

Remark 2.4.2 *In the above scheme the grid in time and space has a mesh size Δ and $\bar{\sigma}\sqrt{\Delta}$ respectively. This ensures that the scheme is monotone (see Lemma 2.4.6). However, when $\bar{\sigma}$ is very large, Δ has to keep very small, thus increasing time steps dramatically. In this case, an implicit finite difference scheme is preferred which is unconditionally monotone with arbitrary mesh sizes.*

Remark 2.4.3 *The above scheme implies that, for any fixed $x \in \mathbb{R}$, $u^{\Delta}(t,x)$ is constant in t between each partition grid in time $n\Delta \leq t < (n+1)\Delta$ for $n \in \mathbb{N}$. In particular, u^{Δ} is equal to the initial condition g inside the first time grid $t < \Delta$.*

Remark 2.4.4 *The algorithm (2.17) implies that $u^\Delta(t,x)$ is equal to a weighted average of $u^\Delta(t-\Delta, x-\bar{\sigma}\sqrt{\Delta})$, $u^\Delta(t-\Delta, x+\bar{\sigma}\sqrt{\Delta})$ and $u^\Delta(t-\Delta, x)$. Further, if we consider random variables $X^{t,x}$ such that $X^{t,x} = \pm\bar{\sigma}$ with probability $\frac{a(t,x)}{2\bar{\sigma}^2}$, and $X^{t,x} = 0$ with probability $1 - \frac{a(t,x)}{\bar{\sigma}^2}$, then we can rewrite (2.17) as*

$$u^\Delta(t,x) = \mathbb{E}[u^\Delta(t-\Delta, x+\sqrt{\Delta}X^{t,x})], \tag{2.18}$$

and correspondingly we have for $t \geq \Delta$,

$$S(\Delta, t, x, r, v) = \frac{r - \mathbb{E}[v(x+\sqrt{\Delta}X^{t,x})]}{\Delta}. \tag{2.19}$$

A comparison property of the scheme S follows immediately, and it will be used later in the convergence analysis.

Lemma 2.4.5 *Suppose that two functions $u, v \in \mathcal{C}_b([0,T]\times\mathbb{R})$ satisfy*

$$S(\Delta, t, x, u, u|_{t-\Delta}) \leq h_1 \quad in\ [\Delta, T]\times\mathbb{R};$$

$$S(\Delta, t, x, v, v|_{t-\Delta}) \geq h_2 \quad in\ [\Delta, T]\times\mathbb{R},$$

for some constants h_1, h_2. Then,

$$u - v \leq \sup_{[0,\Delta)\times\mathbb{R}} (u-v) + T(h_1-h_2)^+ \quad in\ [0,T]\times\mathbb{R}. \tag{2.20}$$

Proof. By writing S as in (2.19), we have

$$u(t,x) \leq \mathbb{E}[u(t-\Delta, x+\sqrt{\Delta}X^{t,x})] + \Delta h_1$$

and

$$v(t,x) \geq \mathbb{E}[v(t-\Delta, x+\sqrt{\Delta}X^{t,x})] + \Delta h_2$$

Combining the above two inequalities then yields

$$(u-v)(t,x) \leq \sup_{y\in\mathbb{R}} (u-v)(t-\Delta, y) + \Delta(h_1-h_2)^+.$$

The conclusion then follows by induction. ■

We now check that the finite difference scheme (2.15) indeed satisfy the *monotonicity*, *stability*, and *consistency* conditions in Barles-Souganidis convergence framework.

Lemma 2.4.6 *(Monotonicity) For any $(\Delta, t, x, r) \in (0,T)\times[0,T]\times\mathbb{R}\times\mathbb{R}$, and $u, v \in \mathcal{C}_b(\mathbb{R})$ such that $u \geq v$,*

$$S(\Delta, t, x, r, u) \leq S(\Delta, t, x, r, v).$$

Proof. It holds obviously when $t < \Delta$. When $t \geq \Delta$, it follows immediately from the assumption that $a(t,x) \geq 0$ and the fact that $1 - \frac{a(t,x)}{\bar{\sigma}^2} \geq 0$. ■

Lemma 2.4.7 *(Stability) For every $\Delta \in (0,T)$, the scheme (2.15) admits a unique bounded solution u^Δ with*

$$|u^\Delta|_0 \leq |g|_0.$$

Proof. The existence and uniqueness of solution u^Δ is trivial. From (2.17), we have that for $t \geq \Delta$,

$$|u^\Delta(t,\cdot)|_0 \leq |u^\Delta(t-\Delta,\cdot)|_0.$$

Thus, by induction, we conclude that $|u^\Delta|_0 \leq |g|_0$. ∎

Lemma 2.4.8 *(Consistency) For any $\varphi \in \mathcal{C}_b^\infty([0,T]\times\mathbb{R})$,*

$$\begin{aligned} &\left| S(\Delta,t,x,\varphi(t,x),\varphi|_{t-\Delta}) - \partial_t\varphi(t,x) + \frac{1}{2}a(t,x)\partial_{xx}\varphi(t,x) \right| \\ &\leq \Delta\left(|\partial_{tt}\varphi|_0 + |a|_0|\partial_{xxt}\varphi|_0 + |a|_0^2|\partial_{xxxx}\varphi|_0\right) \end{aligned} \tag{2.21}$$

in $[\Delta,T]\times\mathbb{R}$.

Proof. From (2.16), we have for $t \geq \Delta$,

$$\begin{aligned} &S(\Delta,t,x,\varphi(t,x),\varphi|_{t-\Delta}) - \partial_t\varphi(t,x) + \frac{1}{2}a(t,x)\partial_{xx}\varphi(t,x) \\ = &\left[\frac{\varphi(t,x)-\varphi(t-\Delta,x)}{\Delta} - \partial_t\varphi(t,x)\right] + \frac{1}{2}a(t,x)\left[\partial_{xx}\varphi(t,x) - \partial_{xx}\varphi(t-\Delta,x)\right] \\ - &\frac{1}{2}a(t,x)\left[\frac{\varphi(t-\Delta,x-\bar\sigma\sqrt{\Delta}) + \varphi(t-\Delta,x+\bar\sigma\sqrt{\Delta}) - 2\varphi(t-\Delta,x)}{\bar\sigma^2\Delta} - \partial_{xx}\varphi(t-\Delta,x)\right] \\ =: &(I) + (II) + (III) \end{aligned}$$

By standard Taylor expansions with integral remainders, we have that for a smooth function f,

$$\begin{aligned} f(x+h) &= f(x) + h\int_0^1 f'(x+sh)ds & (2.22) \\ &= f(x) + hf'(x) + h^2\int_0^1 (1-s)f''(x+sh)ds & (2.23) \\ &= f(x) + hf'(x) + \frac{1}{2}h^2 f''(x) + \frac{1}{6}h^3 f'''(x) \\ &+ \frac{1}{6}h^4\int_0^1 (1-s)^3 f''''(x+sh)ds & (2.24) \end{aligned}$$

Applying (2.22), (2.23) and (2.24) to term (II), (I) and (III) respectively yields

$$|(I)| \leq \Delta|\partial_{tt}\varphi|_0, \quad |(II)| \leq \Delta|a|_0|\partial_{xxt}\varphi|_0, \quad |(III)| \leq \Delta|a|_0^2|\partial_{xxxx}\varphi|_0.$$

The conclusion then follows by combining the above three estimates. ∎

Clearly, (2.21) implies that the consistency condition (2.11)-(2.12) is satisfied, and thus by Theorem 2.3.2, we have proved

$$u^\Delta \to u \quad \text{uniformly in } [0,T]\times\mathbb{R} \;\text{ as } \Delta \to 0.$$

To obtain the rate of convergence of u^Δ to u, we need a uniform error between them which depends only on some power of Δ. We next establish the error estimates for $u - u^\Delta$. A classical idea is to find a sequence of smooth classical solutions $\{u_\varepsilon\}_{\varepsilon>0}$ of (2.13) that approximates u uniformly when ε is small, and plug them into the consistency error estimate (2.21) and thus obtain error estimates of $u_\varepsilon - u^\Delta$ due to some comparison properties for the scheme S. The error estimate for $u - u^\Delta$ then follows immediately by combining the estimates of $u - u_\varepsilon$ and $u_\varepsilon - u^\Delta$, and an optimization with respect to ε.

However, this idea is not always easy to implement in general and we will show later that it is only implementable in our example in the case where the diffusion coefficient $a(t,x)$ is a constant. When $a(t,x)$ is variable but with some regularity properties, however, we can modify the above classical procedure by applying the *shaking the coefficients* technique introduced by Krylov.

In the following, we need some regularity assumption for the diffusion coefficient a and the initial condition g:

Assumption 2.4.9 *$g(x)$ is Lipschitz and $\sigma(t,x) := \sqrt{a(t,x)}$ is Lipschitz in x and $\frac{1}{2}$-Hölder in t: There exists a constant $M > 0$ such that*

$$|g|_1 + |\sigma|_1 \le M$$

2.4.1 Error estimates when $a(t,x)$ is constant

When $a(t,x) \equiv a = \sigma^2$ is a constant, (2.13) reduces to heat equation and it is standard that the solution u has a representation

$$u(t,x) = \mathbb{E}[g(x + \sigma\sqrt{t}Z)] \tag{2.25}$$

where Z is a standard normal random variable. This, together with the boundedness and Lipschitz continuity of g, implies that $u \in \mathcal{C}_b^1([0,T] \times \mathbb{R})$, with

$$|u|_1 \le C \tag{2.26}$$

for some constant C depending only on M.

We can then easily find a sequence of smooth solutions $\{u_\varepsilon\}_{\varepsilon>0}$ of (2.13) which approximates u by a standard regularization by convolution procedure as follows:

Let $\rho(t,x)$ be a nonnegative smooth function with support in $(-1,0)\times(-1,1)$ and mass 1, and introduce the sequence of mollifiers ρ_ε for small $\varepsilon > 0$,

$$\rho_\varepsilon(t,x) := \frac{1}{\varepsilon^3}\rho\left(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\right). \tag{2.27}$$

For $(t,x) \in [0,T] \times \mathbb{R}$, we define

$$u_\varepsilon(t,x) = u * \rho_\varepsilon(t,x) = \int_{-\varepsilon^2<\tau<0} \int_{|e|<\varepsilon} u(t-\tau, x-e)\rho_\varepsilon(\tau,e)de d\tau,$$

where we extend the domain of solution u to $[0, T+\varepsilon^2] \times \mathbb{R}$. We then have

Lemma 2.4.10 *For any small $\varepsilon > 0$, u_ε is a (viscosity) solution of (2.13).*

Proof. Since $a(t,x)$ is constant, we have for any $(\tau,e) \in (-\varepsilon^2, 0) \times (-\varepsilon, \varepsilon)$, the function $v^{(\tau,e)}(t,x) := u(t-\tau, x-e)$ satisfies the equation (2.13). On the other hand, a Riemann sum approximation shows that there exists a sequence of functions $\{I_n\}_{n\geq 1}$ on $[0,T]\times\mathbb{R}$ such that each I_n is a convex combination of the functions $v^{(\tau,e)}$ for different (τ,e), and that I_n converges uniformly to u_ε.

Since each $v^{(\tau,e)}$ is a solution of the linear equation (2.13), we know I_n is still a solution of (2.13). The conclusion then follows by applying the stability results of viscosity solutions (see Proposition 2.2.6). ■

On the other hand, due to the regularity result (2.26) of u, standard properties of mollifiers imply that $u_\varepsilon \in \mathcal{C}_b^\infty([0,T]\times\mathbb{R})$,

$$|u-u_\varepsilon|_0 \leq \int_{-\varepsilon^2<\tau<0}\int_{|e|<\varepsilon} |u(t,x)-u(t-\tau,x-e)|_0\,\rho_\varepsilon(\tau,e)ded\tau \leq C\varepsilon, \tag{2.28}$$

and, moreover, for positive integer i and j, we have by change of variable that

$$\partial_t^i\partial_x^j u_\varepsilon(t,x) = \varepsilon^{-2i-j}\int_{-1<s<0}\int_{|z|<1} u(t-\varepsilon^2 s, x-\varepsilon z)\partial_t^i\partial_x^j\rho(s,z)dzds.$$

Since for $i\neq 0$ or $j\neq 0$,

$$\int_{-1<s<0}\int_{|z|<1}\partial_t^i\partial_x^j\rho(s,z)dzds = 0,$$

we further obtain that

$$\begin{aligned}|\partial_t^i\partial_x^j u_\varepsilon|_0 &\leq \varepsilon^{-2i-j}\int_{-1<s<0}\int_{|z|<1}\left|u(t-\varepsilon^2 s,x-\varepsilon z)-u(t,x)\right|_0\left|\partial_t^i\partial_x^j\rho(s,z)\right|dzds\\ &\leq C\varepsilon^{1-2i-j},\end{aligned} \tag{2.29}$$

where the constant C depends only on M.

We are now able to plug u_ε into the consistency error estimate (2.21), which gives us

$$|S(\Delta,t,x,u_\varepsilon(t,x),u_\varepsilon|_{t-\Delta})| \leq \Delta\left(|\partial_{tt}u_\varepsilon|_0 + a|\partial_{xxt}u_\varepsilon|_0 + a^2|\partial_{xxxx}u_\varepsilon|_0\right) \leq C\Delta\varepsilon^{-3}$$

in $[\Delta,T]\times\mathbb{R}$, where the last inequality follows by the estimates (2.29) and the constant C depends only on M.

Comparing u_ε and u^Δ by applying the comparison property (2.20) then yields

$$|u_\varepsilon-u^\Delta| \leq \sup_{[0,\Delta)\times\mathbb{R}}|u_\varepsilon-u^\Delta| + CT\Delta\varepsilon^{-3}\ \text{ in } [0,T]\times\mathbb{R}.$$

Finally, by (2.26) and (2.28), we conclude

$$\begin{aligned}|u-u^\Delta| &\leq |u-u_\varepsilon| + |u_\varepsilon-u^\Delta|\\ &\leq \sup_{[0,\Delta)\times\mathbb{R}}|u-u^\Delta| + CT\Delta\varepsilon^{-3}+C\varepsilon\\ &= \sup_{[0,\Delta)\times\mathbb{R}}|u-g| + CT\Delta\varepsilon^{-3}+C\varepsilon\\ &\leq C\sqrt{\Delta}+CT\Delta\varepsilon^{-3}+C\varepsilon \leq C\Delta^{\frac{1}{4}}\ \text{ in } [0,T]\times\mathbb{R},\end{aligned}$$

by choosing $\varepsilon=\Delta^{\frac{1}{4}}$ in the last inequality, where the constant C depends only on M and T.

2.4.2 Error estimates when $a(t,x)$ is variable

Unfortunately, when $a(t,x)$ is variable, Lemma 2.4.10 fails to hold since in this case the functions $v^{(\tau,e)}(t,x) := u(t-\tau, x-e)$ do not satisfy the equation (2.13). Instead, each $v^{(\tau,e)}(t,x)$ is a solution of (2.13) with the diffusion coefficient $a(t,x)$ perturbed to $a(t-\tau, x-e)$. Inspired by this, Krylov [43, 44] proposed the *shaking the coefficients* technique to build a sequence of smooth subsolutions (or supersolutions) instead of smooth solutions. This will provide a one-sided error estimate, namely upper bound (or lower bound) of $u - u^{\Delta}$.

Thanks to the linearity of equation (2.13), we are able to apply this technique twice but with opposite directions to build both a sequence of smooth subsolutions and a sequence of smooth supersolutions, and thus obtain both upper and lower bound of $u - u^{\Delta}$.

To this end, for small $\varepsilon > 0$, we extend the diffusion coefficient $a(t,x)$ to $(-\varepsilon^2, T+\varepsilon^2] \times \mathbb{R}$ such that Assumption 2.4.9 still holds, and first consider a perturbed equation

$$\partial_t u^\varepsilon - \inf_{\tau \in (-\varepsilon^2, 0), |e| \le \varepsilon} \left\{ \frac{1}{2} a(t+\tau, x+e) \partial_{xx} u^\varepsilon \right\} = 0 \ \text{ in } (0, T+\varepsilon^2] \times \mathbb{R}, \quad (2.30)$$

with initial condition

$$u^\varepsilon(0,x) = g(x).$$

We have existence, uniqueness, and regularity results for equation (2.30) and a comparison between u and u^ε. A similar result for a more general type of equations will be presented in the following chapters of this thesis with its proof in detail, thus we skip the proof here.

Proposition 2.4.11 *Suppose that Assumption 2.4.9 holds. Then there exists a unique viscosity solution $u^\varepsilon \in \mathcal{C}_b^1([0, T+\varepsilon^2] \times \mathbb{R})$ of the equation (2.30), with $|u^\varepsilon|_1 \le C$, for some constant C depending only on M and T. Moreover,*

$$|u - u^\varepsilon| \le C\varepsilon \quad \text{in } [0,T] \times \mathbb{R} \quad (2.31)$$

We then follow a similar regularization procedure by defining for each $(t,x) \in [0,T] \times \mathbb{R}$,

$$u_\varepsilon(t,x) = u^\varepsilon * \rho_\varepsilon(t,x) = \int_{-\varepsilon^2 < \tau < 0} \int_{|e| < \varepsilon} u^\varepsilon(t-\tau, x-e) \rho_\varepsilon(\tau, e) de d\tau,$$

which implies again that $u_\varepsilon \in \mathcal{C}_b^\infty([0,T] \times \mathbb{R})$,

$$|u^\varepsilon - u_\varepsilon|_0 \le C\varepsilon, \quad (2.32)$$

and for positive integer i and j,

$$|\partial_t^i \partial_x^j u_\varepsilon|_0 \le C\varepsilon^{1-2i-j}, \quad (2.33)$$

where the constant C depends only on M.

Notice that by (2.30), we have that for any $(\tau, e) \in (-\varepsilon^2, 0) \times (-\varepsilon, \varepsilon)$, the function $v^{(\tau,e)}(t,x) := u(t-\tau, x-e)$ becomes a viscosity subsolution of the equation (2.13). Similar to Lemma 2.4.10, we now have

Lemma 2.4.12 *For any small $\varepsilon > 0$, u_ε is a (viscosity) subsolution of (2.13).*

Now we plug u_ε into the consistency error estimate (2.21) and obtain a one-sided estimate

$$S(\Delta, t, x, u_\varepsilon(t,x), u_\varepsilon|_{t-\Delta}) \leq \Delta\left(|\partial_{tt}u_\varepsilon|_0 + |a|_0|\partial_{xxt}u_\varepsilon|_0 + |a|_0^2|\partial_{xxxx}u_\varepsilon|_0\right) \leq C\Delta\varepsilon^{-3}$$

in $[\Delta, T] \times \mathbb{R}$, where the constant C depends only on M.

Comparing u_ε and u^Δ by applying the comparison property (2.20) now yields

$$u_\varepsilon - u^\Delta \leq \sup_{[0,\Delta)\times\mathbb{R}} (u_\varepsilon - u^\Delta) + CT\Delta\varepsilon^{-3} \quad \text{in } [0,T] \times \mathbb{R}.$$

Finally, by (2.31) and (2.32), we have

$$\begin{aligned}
u - u^\Delta &= (u - u^\varepsilon) + (u^\varepsilon - u_\varepsilon) + (u_\varepsilon - u^\Delta) \\
&\leq \sup_{[0,\Delta)\times\mathbb{R}} |u - u^\Delta| + CT\Delta\varepsilon^{-3} + C\varepsilon \\
&= \sup_{[0,\Delta)\times\mathbb{R}} |u - g| + CT\Delta\varepsilon^{-3} + C\varepsilon \\
&\leq C\sqrt{\Delta} + CT\Delta\varepsilon^{-3} + C\varepsilon \leq C\Delta^{\frac{1}{4}} \quad \text{in } [0,T] \times \mathbb{R},
\end{aligned}$$

by choosing $\varepsilon = \Delta^{\frac{1}{4}}$ in the last inequality, where the constant C depends only on M and T.

We have now derived the upper bound of $u - u^\Delta$. To get the lower bound, we only need to consider another perturbed equation which is in the opposite direction of equation (2.30):

$$\partial_t u^\varepsilon - \sup_{\tau\in(-\varepsilon^2,0),|e|\leq 1} \left\{\frac{1}{2}a(t+\tau, x+e)\partial_{xx}u^\varepsilon\right\} = 0 \quad \text{in } (0, T+\varepsilon^2] \times \mathbb{R}, \quad (2.34)$$

and follow the same regularization procedure as before. Thus, we have proved

Theorem 2.4.13 *Suppose that Assumption 2.4.9 holds. Then, there exists a constant C depending only on M and T such that*

$$|u - u^\Delta| \leq C\Delta^{\frac{1}{4}} \quad \textit{in } [0,T] \times \mathbb{R}.$$

Remark 2.4.14 *A vital reason that we are able to reverse the direction of perturbation as in equation (2.34) and obtain a similar lemma to Lemma 2.4.12 but in the opposite direction (i.e. supersolution instead of subsolution) is that equation (2.13) is linear in the sense that it is both convex and concave in the terms involving the derivatives of u. When the equation is either convex or concave only, we can only perturb the equation in one direction and thus obtain only a one-sided error estimate. To obtain the error estimate for the other side, two techniques that are commonly used in the literature are interchanging the roles of schemes and equations, given enough regularity of the scheme solutions, and optimal switching approximations. Both of the two techniques will be used later in this thesis.*

2.5 Convergence rate of classical Central Limit Theorem (CLT)

Let $\{X_n : n \geq 1\}$ be a sequence of i.i.d random variables with mean 0 and variance 1, and define $\hat{S}_n := n^{-\frac{1}{2}} \sum_1^n X_k$. The standard central limit theorem shows that $\hat{S}_n$ converges to standard Gaussian in distribution, namely

$$\lim_{n\to\infty} \mathbb{E}[g(\hat{S}_n)] = \mathbb{E}[g(Z)], \tag{2.35}$$

where Z is standard Gaussian random variable and g is any continuous function satisfying quadratic growth condition.

In this section, we apply the convergence analysis introduced in the last section to obtain a convergence rate of CLT. To this end, we consider again equation (2.13)-(2.14) with $a(t,x) \equiv 1$. Then, from (2.25) we have

$$u(1,0) = \mathbb{E}[g(Z)], \tag{2.36}$$

which is the right hand side of (2.35). For the left hand side, if we are able to construct an approximation u^Δ of u such that

$$u^{\frac{1}{n}}(1,0) = \mathbb{E}[g(\hat{S}_n)], \tag{2.37}$$

then we directly obtain the convergence rate using the error estimates of $u - u^\Delta$. Indeed, if we introduce $u^\Delta : [0,1] \times \mathbb{R} \to \mathbb{R}$ recursively as

$$u^\Delta(t,x) = \mathbb{E}[u^\Delta(t-\Delta, x + \sqrt{\Delta}X^t)]\mathbf{1}_{\{t\geq\Delta\}} + g(x)\mathbf{1}_{\{t<\Delta\}}, \tag{2.38}$$

where X^t is X_n for $n\Delta \leq t < (n+1)\Delta$, then we immediately have (2.37) by letting $\Delta = 1/n$ and $x = 0$ in the following lemma:

Lemma 2.5.1 *For $x \in \mathbb{R}$ and $n \in \mathbb{N}$ such that $n\Delta \leq 1$,*

$$u^\Delta(n\Delta, x) = \mathbb{E}[g(x + \sqrt{\Delta}\sum_{k=1}^n X_k)]. \tag{2.39}$$

Proof. We prove by induction on n. For $n = 1$, we have by (2.38) that

$$u^\Delta(\Delta, x) = \mathbb{E}[u^\Delta(0, x + \sqrt{\Delta}X_1)] = \mathbb{E}[g(x + \sqrt{\Delta}X_1)].$$

Suppose (2.39) holds for some $n \in \mathbb{N}$, then

$$\begin{aligned} u^\Delta((n+1)\Delta, x) &= \mathbb{E}[u^\Delta(n\Delta, x + \sqrt{\Delta}X_{n+1})] \\ &= \mathbb{E}\left[\mathbb{E}\left[g(x + \sqrt{\Delta}y + \sqrt{\Delta}\sum_{k=1}^n X_k)\right]_{y=X_{n+1}}\right] \\ &= \mathbb{E}[g(x + \sqrt{\Delta}\sum_{k=1}^{n+1} X_k)], \end{aligned}$$

where the last equality follows from the mutually independence of $\{X_k\}$. ■

By (2.36) and (2.37), we are left to establish the error estimate of $u(1,0)$

and $u^{\frac{1}{n}}(1,0)$. To this end, we write the approximation scheme (2.38) as

$$S(\Delta, t, x, u^{\Delta}(t,x), u^{\Delta}|_{t-\Delta}) = 0 \quad \text{in } [0,1]\times\mathbb{R} \tag{2.40}$$

where

$$S(\Delta, t, x, r, v) := \begin{cases} \frac{r - \mathbb{E}[v(x+\sqrt{\Delta}X^t)]}{\Delta} & \text{if } t \geq \Delta \\ r - g(x) & \text{if } t < \Delta \end{cases} \tag{2.41}$$

Following the same argument as in the proof of Lemma 2.4.5, we have that the comparison result of the scheme (2.40) still holds. In particular, we have:

Lemma 2.5.2 *Suppose that two functions $u, v \in \mathcal{C}_b([0,1]\times\mathbb{R})$ satisfy*

$$S(\Delta, t, x, u, u|_{t-\Delta}) \leq h_1 \quad \textit{in } [\Delta, 1]\times\mathbb{R};$$

$$S(\Delta, t, x, v, v|_{t-\Delta}) \geq h_2 \quad \textit{in } [\Delta, 1]\times\mathbb{R},$$

for some constants h_1, h_2. Then, for $x \in \mathbb{R}$ and $n \in \mathbb{N}$ such that $n\Delta \leq 1$,

$$(u-v)(n\Delta, x) \leq \sup_{\{0\}\times\mathbb{R}} (u-v) + (h_1 - h_2)^+. \tag{2.42}$$

However, the consistency error estimate has now changed from (2.21). We now impose some assumptions on g and the random variable X_n, and establish two versions of consistency error estimate depending on the regularity of the test function φ.

Assumption 2.5.3 *(i) $g : \mathbb{R} \to \mathbb{R}$ is bounded and Lipschitz continuous with $|g|_1 \leq M$ for some constant $M > 0$.*

(ii) The sequence of i.i.d. random variables X_n has a finite third moment: $M_X^3 < \infty$, with $M_X^p := \mathbb{E}[|X_1|^p]$ for $p > 0$. Moreover, X_n has mean 0 and variance 1, i.e. $\mathbb{E}[X_n] = 0$ and $\mathbb{E}[X_n^2] = 1$.

Lemma 2.5.4 *Suppose that Assumption 2.5.3(ii) is satisfied. Then, the following consistency error estimates holds for the scheme (2.40):*

(a) If $\varphi \in \mathcal{C}_b^\infty([0,1]\times\mathbb{R})$, then

$$\left| S(\Delta, t, x, \varphi(t,x), \varphi|_{t-\Delta}) - \partial_t\varphi(t-\Delta, x) + \frac{1}{2}\partial_{xx}\varphi(t-\Delta, x) \right|$$
$$\leq \Delta|\partial_{tt}\varphi|_0 + \sqrt{\Delta}M_X^3|\partial_{xxx}\varphi|_0 \quad \textit{in } [\Delta,1]\times\mathbb{R}. \tag{2.43}$$

(b) if $\varphi \in \mathcal{C}_b^{1+\frac{1}{2},2+1}([0,1]\times\mathbb{R})$, then

$$\left| S(\Delta, t, x, \varphi(t,x), \varphi|_{t-\Delta}) - \partial_t\varphi(t-\Delta, x) + \frac{1}{2}\partial_{xx}\varphi(t-\Delta, x) \right|$$
$$\leq \sqrt{\Delta}\left([\partial_t\varphi]_{\mathcal{C}^{1/2,1}} + M_X^3[\partial_{xx}\varphi]_{\mathcal{C}^{1/2,1}}\right) \quad \textit{in } [\Delta,1]\times\mathbb{R}. \tag{2.44}$$

Proof. For $t \geq \Delta$ and $x \in \mathbb{R}$,

$$
\begin{aligned}
& \left| S(\Delta, t, x, \varphi(t,x), \varphi|_{t-\Delta}) - \partial_t \varphi(t-\Delta, x) + \frac{1}{2}\partial_{xx}\varphi(t-\Delta, x) \right| \\
\leq\ & |\varphi(t,x) - \varphi(t-\Delta, x) - \Delta \partial_t \varphi(t-\Delta, x)|\, \Delta^{-1} \\
+\ & \left| \mathbb{E}[\varphi(t-\Delta, x+\sqrt{\Delta}X^t) - \varphi(t-\Delta, x)] - \frac{1}{2}\Delta \partial_{xx}\varphi(t-\Delta, x) \right| \Delta^{-1} \\
\leq\ & \Delta^{-1} \int_{t-\Delta}^{t} |\partial_t \varphi(s,x) - \partial_t \varphi(t-\Delta, x)|\, ds \\
+\ & \Delta^{-1}\mathbb{E}\left[\int_x^{x+\sqrt{\Delta}X^t} \int_x^s |\partial_{xx}\varphi(t-\Delta, p) - \partial_{xx}\varphi(t-\Delta, x)| dpds \right] \\
:=\ & (I) + (II),
\end{aligned}
$$

where we used the assumptions that $\mathbb{E}[X^t] = 0$ and $\mathbb{E}[|X^t|^2] = 1$.

Then, in case (a), we have

$$
(I) \leq \Delta^{-1} \int_{t-\Delta}^{t} |\partial_{tt}\varphi|_0 |s - t + \Delta| ds = \frac{1}{2}\Delta |\partial_{tt}\varphi|_0,
$$

$$
(II) \leq \Delta^{-1}\mathbb{E}\left[\int_x^{x+\sqrt{\Delta}X^t} \int_x^s |\partial_{xxx}\varphi|_0 |p - x| dpds \right] = \frac{1}{6}\sqrt{\Delta} M_X^3 |\partial_{xxx}\varphi|_0.
$$

Similarly, in case (b),

$$
(I) \leq \Delta^{-1} \int_{t-\Delta}^{t} [\partial_t \varphi]_{\mathcal{C}^{1/2,1}} |s - t + \Delta|^{\frac{1}{2}} ds = \frac{2}{3}\sqrt{\Delta}[\partial_t \varphi]_{\mathcal{C}^{1/2,1}},
$$

$$
(II) \leq \Delta^{-1}\mathbb{E}\left[\int_x^{x+\sqrt{\Delta}X^t} \int_x^s [\partial_{xx}\varphi]_{\mathcal{C}^{1/2,1}} |p - x| dpds \right] = \frac{1}{6}\sqrt{\Delta} M_X^3 [\partial_{xx}\varphi]_{\mathcal{C}^{1/2,1}}.
$$

We then conclude by combining the above estimates. ■

We now give the convergence rate of classical central limit theorem (CLT) in Theorem 2.5.5. The convergence result (2.46) has already existed in the literature (see e.g. [61]) with a similar assumption that $g \in \mathcal{C}_b^3(\mathbb{R})$, but here we apply a novel approach by building a connection between convergence of numerical schemes and CLT, and only require that $g \in \mathcal{C}_b^{2+1}(\mathbb{R})$. Moreover, by utilizing a mollification procedure, we can also obtain a more general convergence result (2.45) when g is only bounded and Lipschitz continuous, and to the best of our knowledge, this is a new result to the literature.

Theorem 2.5.5 *Suppose that Assumption 2.5.3 is satisfied. Then, there exists a constant C depending only on M and M_X^3 such that*

$$
\left| \mathbb{E}[g(\hat{S}_n)] - \mathbb{E}[g(Z)] \right| \leq C n^{-\frac{1}{6}}. \tag{2.45}
$$

Moreover, if $g \in \mathcal{C}_b^{2+1}(\mathbb{R})$, then

$$
\left| \mathbb{E}[g(\hat{S}_n)] - \mathbb{E}[g(Z)] \right| \leq (1 + M_X^3)[g'']_{\mathcal{C}^1} n^{-\frac{1}{2}}. \tag{2.46}
$$

Proof. By (2.36) and (2.37), it suffices to establish the error estimates of $(u-u^{\frac{1}{n}})(1,0)$. For this, we apply again the regularization procedure introduced in Section 2.4.1: define $u_\varepsilon(t,x) := u * \rho_\varepsilon(t,x)$ and we have $|u-u_\varepsilon|_0 \le C\varepsilon$, $|\partial_t^i\partial_x^j u_\varepsilon|_0 \le C\varepsilon^{1-2i-j}$, and by (2.43),

$$|S(\Delta,t,x,u_\varepsilon(t,x),u_\varepsilon|_{t-\Delta})| \le \Delta|\partial_{tt}u_\varepsilon|_0+\sqrt{\Delta}M_X^3|\partial_{xxx}u_\varepsilon|_0 \le C(\Delta\varepsilon^{-3}+\sqrt{\Delta}\varepsilon^{-2})$$

in $[\Delta,1]\times\mathbb{R}$, where the constant C depends only on M and M_X^3. Then from (2.42) we obtain that

$$(u_\varepsilon - u^{\frac{1}{n}})(1,0) \le \sup_{\{0\}\times\mathbb{R}}(u_\varepsilon - u^{\frac{1}{n}}) + C(n^{-1}\varepsilon^{-3}+n^{-1/2}\varepsilon^{-2}),$$

and thus,

$$\begin{aligned}|u-u^{\frac{1}{n}}|(1,0) &\le |u-u_\varepsilon|(1,0)+|u_\varepsilon-u^{\frac{1}{n}}|(1,0)\\ &\le C(n^{-1}\varepsilon^{-3}+n^{-1/2}\varepsilon^{-2})+C\varepsilon \le Cn^{-\frac{1}{6}} \text{ in } [0,1]\times\mathbb{R},\end{aligned}$$

by choosing $\varepsilon = n^{-\frac{1}{6}}$ in the last inequality, where the constant C depends only on M and M_X^3.

Moreover, if $g\in\mathcal{C}_b^{2+1}(\mathbb{R})$, then since $u(t,x)=\mathbb{E}[g(x+\sqrt{t}Z)]$, we have $u\in\mathcal{C}_b^{1+\frac{1}{2},2+1}([0,1]\times\mathbb{R})$ with

$$[\partial_t u]_{\mathcal{C}^{1/2,1}} = \frac{1}{2}[\partial_{xx}u]_{\mathcal{C}^{1/2,1}} \le \frac{1}{2}[g'']_{\mathcal{C}^1}.$$

Plugging u into (2.44) then yields

$$|S(\Delta,t,x,u(t,x),u|_{t-\Delta})| \le (1+M_X^3)[g'']_{\mathcal{C}^1}\sqrt{\Delta}$$

in $[\Delta,1]\times\mathbb{R}$. Again from (2.42) we obtain directly that

$$|u-u^{\frac{1}{n}}|(1,0) \le (1+M_X^3)[g'']_{\mathcal{C}^1}n^{-\frac{1}{2}}.$$

■

Remark 2.5.6 *We can obtain the same convergence rate of $\frac{1}{4}$ as in Theorem 2.4.13, if the random variables X_n satisfies further that $\mathbb{E}[X_n^3]=0$ and $M_X^4=\mathbb{E}[|X_n|^4]<\infty$ (which are satisfied by the random variable $X^{t,x}$ specified in Remark 2.4.4). Indeed, in this case the right hand side of inequality (2.43) becomes $\Delta(|\partial_{tt}\varphi|_0+M_X^4|\partial_{xxxx}\varphi|_0)$ and thus improves the final rate of convergence.*

2.5.1 Berry-Esseen type of convergence

In this section, we are interested in the rate at which the distribution functions

$$F_n(x) := \mathbb{P}(\hat{S}_n \le x) = \mathbb{E}\left[\mathbf{1}_{\{\hat{S}_n\le x\}}\right]$$

converges to the Gaussian distribution function

$$G(x) := \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \mathbb{E}\left[\mathbf{1}_{\{Z \le x\}}\right].$$

This work was firstly done independently by Berry [7] and Esseen [24] in 1940s, where the so-called Berry-Esseen theorem states that there exists a constant C such that

$$|F_n - G|_0 \le C M_X^3 n^{-\frac{1}{2}}. \tag{2.47}$$

Following that, Berry and Esseen, along with other mathematicians, refined the constant C repeatedly over the subsequent decades.

Here we are not concerned with how large the value of constant C is but try to establish a similar convergence result to (2.47) with an explicit constant C by applying the estimate (2.46). Unfortunately, we can not directly choose $g(y) = \mathbf{1}_{\{y \le x\}}$ in (2.46) since the step function is not continuous and differentiable at x. However, note that using integration by parts, we have for $g \in \mathcal{C}_b^1(\mathbb{R})$,

$$\begin{aligned} & \int g'(x)(G(x) - F_n(x))dx \\ = & \int g(x)dF_n(x) - \int g(x)dG(x) \\ = & \mathbb{E}[g(\hat{S}_n)] - \mathbb{E}[g(Z)]. \end{aligned} \tag{2.48}$$

We may then choose some $g \in \mathcal{C}_b^{2+1}(\mathbb{R})$ to apply (2.46) in the right hand side of (3.29) while at the same time getting some information about $|F_n - G|$ in the left hand side. Indeed, following this idea we find a much simpler way to establish a rate of Berry-Esseen type convergence but with a slower rate of 1/8.

Proposition 2.5.7 *Suppose that Assumption 2.5.3(ii) is satisfied. Then,*

$$|F_n - G|_0 \le (M_X^3 + 5) n^{-\frac{1}{8}}. \tag{2.49}$$

Proof. Let $h : \mathbb{R} \to \mathbb{R}$ be such that $h(0) = 0$ and

$$h'(x) = x^2 \mathbf{1}_{\{0 \le x \le 1\}} + (2 - (x-2)^2)\mathbf{1}_{\{1 < x \le 3\}} + (x-4)^2 \mathbf{1}_{\{3 < x \le 4\}},$$

then it is obvious that $h \in \mathcal{C}_b^{2+1}(\mathbb{R})$ with $[h'']_1 = 2$. Define for any $a \in \mathbb{R}$ and $\varepsilon > 0$,

$$g_{a,\varepsilon}(x) := h\left(\frac{x-a}{\varepsilon}\right),$$

we then have $g_{a,\varepsilon} \in \mathcal{C}_b^{2+1}(\mathbb{R})$ with $[g''_{a,\varepsilon}]_1 = 2\varepsilon^{-3}$. Applying $g_{a,\varepsilon}$ to (2.46) and (3.29) yields

$$\left| \int g'_{a,\varepsilon}(x)(G(x) - F_n(x))dx \right| \le 2(M_X^3 + 1)\varepsilon^{-3} n^{-\frac{1}{2}}. \tag{2.50}$$

Focussing on the integral of the above left hand side, we further have that by change of variable $y = \frac{x-a}{\varepsilon}$,

$$\begin{aligned}
&\int g'_{a,\varepsilon}(x)(G(x) - F_n(x))dx \\
= &\int h'(y)(G(a+\varepsilon y) - F_n(a+\varepsilon y))dy \\
\leq &\int h'(y)(G(a) + \varepsilon y - F_n(a)dy \\
\leq &\ (G(a) - F_n(a) + 4\varepsilon)\int_0^4 h'(y)dy \\
= &\ 4(G(a) - F_n(a) + 4\varepsilon) \qquad (2.51)
\end{aligned}$$

where we use that $h' \geq 0$, $[G]_1 \leq 1$ and that F_n is monotone. Combining (2.50) and (2.51), we have

$$F_n(a) - G(a) \leq \frac{1}{2}(M_X^3 + 1)\varepsilon^{-3}n^{-\frac{1}{2}} + 4\varepsilon \leq (M_X^3 + 5)n^{-\frac{1}{8}}$$

by letting $\varepsilon = n^{-1/8}$. Similarly,

$$\begin{aligned}
&\int h'(y)(G(a+\varepsilon y) - F_n(a+\varepsilon y))dy \\
\geq &\ (G(a+4\varepsilon) - F_n(a+4\varepsilon) - 4\varepsilon)\int_0^4 h'(y)dy \\
= &\ 4(G(a+4\varepsilon) - F_n(a+4\varepsilon) - 4\varepsilon),
\end{aligned}$$

and thus

$$G(a + 4n^{-1/8}) - F_n(a + 4n^{-1/8}) \leq (M_X^3 + 5)n^{-\frac{1}{8}}$$

By the arbitrary of a, we conclude that

$$|F_n - G|_0 \leq (M_X^3 + 5)n^{-\frac{1}{8}}.$$

■

Remark 2.5.8 *Although the rate of 1/8 we established is slower than that of 1/2 in the Berry-Esseen theorem due to the introduction of an approximation to step functions by $\mathcal{C}^{2+1}$ functions ($g_{a,\varepsilon}$ in the proof above), we showed here a much simpler way to establish a rate of Berry-Esseen type convergence in CLT.*

Chapter 3

A monotone approximation scheme for semilinear parabolic PDEs with convex and coercive Hamiltonians

In this chapter, we propose a monotone approximation scheme for a class of semilinear parabolic equations that are convex and coercive in their gradients. Such equations arise often in pricing and portfolio management in incomplete markets and, more broadly, are directly connected to the representation of solutions to backward stochastic differential equations. The proposed scheme is based on splitting the equation in two parts, the first corresponding to a linear parabolic equation and the second to a Hamilton-Jacobi equation. The solutions of these two equations are approximated using, respectively, the Feynman-Kac and the Hopf-Lax formulae. We then establish the convergence analysis of the scheme by applying Krylov's shaking coefficients technique. However, as mentioned by the last chapter (see Remark 2.4.14), since the equation is not linear, we are only able to obtain a one-sided error estimate from this technique. To obtain the other side's estimate, we will use an optimal switching approximation method introduced by Barles and Jacobsen [3, 4] in Section 3.3.2.

3.1 Introduction

We consider d-dimensional semilinear parabolic equations of the form

$$\begin{cases} -\partial_t u - \dfrac{1}{2}\mathrm{tr}\left(\sigma\sigma^T(t,x)D_x^2 u\right) - b(t,x)\cdot D_x u + H(t,x,D_x u) = 0 & \text{in } Q_T; \\ u(T,x) = U(x) & \text{in } \mathbb{R}^d, \end{cases} \tag{3.1}$$

where $Q_T = [0,T) \times \mathbb{R}^d$, σ is a $d \times d$ matrix, b is a $\mathbb{R}^d$-valued vector. A key feature is that the nonlinear Hamiltonian $H(t,x,p)$ is convex and coercive in p, the latter meaning $H(t,x,p)$ has superlinear grows in infinity with respect to p. In particular, the coercivity covers the case that H has quadratic growth in p, a case that corresponds to a rich class of equations in mathematical finance arising, for example, in optimal investment with homothetic risk preferences

([35]), exponential indifference valuation ([32, 33]) and entropic risk measures ([15]), just to name a few.

More broadly, these equations are inherently connected to (quadratic) backward stochastic differential equations (BSDE), a central area of stochastic analysis ([22, 40, 59]). Specifically, the Hamiltonian $H(t,x,p)$ is directly related to the BSDE's driver and, moreover, the solution of (3.1) yields a functional-form representation of the processes solving the BSDE.

General existence and uniqueness results can be found, among others in [40] as well as in [35], where BSDE techniques have been mainly applied. Closed-form solutions can be constructed only in one-dimensional cases ([67]). Furthermore, approximation schemes have been developed; see [10] and [14] for more references.

In this chapter, we contribute to further studying problem (3.1) by proposing a new approximation scheme. The key idea is to use in an essential way the *convexity* of the Hamiltonian with respect to the gradient. This property is natural in all above applications but it has not been adequately exploited in the existing approximation studies.

To highlight the main ideas and build intuition, we start with some preliminary informal arguments, considering for simplicity slightly simpler equations. To this end, consider the Hamilton-Jacobi (HJ) equation

$$\begin{cases} -\partial_t u + H(D_x u) = 0 & \text{in } Q_T; \\ u(T,x) = U(x) & \text{in } \mathbb{R}^d, \end{cases} \tag{3.2}$$

where the Hamiltonian H is convex and coercive, and the terminal datum U is bounded and Lipschitz continuous. Let L be the Legendre (convex dual) transform of H, $L(q) = \sup_{p\in\mathbb{R}^d}\{p\cdot q - H(p)\}$. The Fenchel-Moreau theorem then yields that $H(p) = \sup_{q\in\mathbb{R}^d}\{p\cdot q - L(q)\}$ and, thus, the HJ equation in (3.2) can be alternatively written as

$$-\partial_t u + \sup_{q\in\mathbb{R}^d}\{D_x u\cdot q - L(D_x u)\} = 0.$$

Classical arguments from control theory then imply the deterministic optimal control representation

$$u(t,x) = \inf_{q\in\mathbb{L}^2[t,T]}\left[\int_t^T L(q_s)ds + U(X_T^{t,x;q})\right],$$

with the controlled state equation $X_s^{t,x;q} = x - \int_t^s q_u du$, for $s\in[t,T]$.

Hopf and Lax observed that, instead of considering the controls in $\mathbb{L}^2[t,T]$, it suffices to optimize over the controls generating geodesic paths of $X^{t,x;q}$, i.e. the constant controls $\hat{q}$ such that $X_T^{t,x;\hat{q}} = y$, for any $y\in\mathbb{R}^d$. Such controls are given by $\hat{q}_s = \frac{x-y}{T-t}$, for $s\in[t,T]$. The above "infinite dimensional" optimal control problem is thus reduced to the "finite dimensional" minimization problem

$$u(t,x) = \inf_{y\in\mathbb{R}^d}\left\{(T-t)L(\frac{x-y}{T-t}) + U(y)\right\}. \quad \text{(Hopf-Lax formula)} \tag{3.3}$$

Adding a diffusion term to equation (3.2) yields the semilinear parabolic equation

$$\begin{cases} -\partial_t u - \frac{1}{2}\mathrm{tr}\left(\sigma\sigma^T(t,x)D_x^2 u\right) + H(D_x u) = 0 & \text{in } Q_T; \\ u(T,x) = U(x) & \text{in } \mathbb{R}^d. \end{cases} \tag{3.4}$$

In analogy to the deterministic case, classical arguments from control theory imply the stochastic optimal control representation

$$u(t,x) = \inf_{q\in\mathbb{H}^2[t,T]} \mathbb{E}\left[\int_t^T L(q_s)ds + U(X_T^{t,x;q})\middle|\, \mathcal{F}_t\right],$$

with the controlled state equation $X_s^{t,x;q} = x - \int_t^s q_u du + \int_t^s \sigma(u, X_u^{t,x;q})dW_u$, for $s\in[t,T]$, and $\mathbb{H}^2[t,T]$ being the space of square-integrable progressively measurable processes q.

Naturally, due to the stochasticity of the state $X^{t,x;q}$, the Hopf-Lax formula (3.3) does not hold for the solution of problem (3.4). On the other hand, note that if we still choose, as in the deterministic case, controls of the form $\hat{q}_s = \frac{x-y}{T-t}$, for $y\in\mathbb{R}^d$ and $s\in[t,T]$, then

$$X_s^{t,x;\hat{q}} = \frac{T-s}{T-t}x + \frac{s-t}{T-t}y + \int_t^s \sigma(u, X_u^{t,x;\hat{q}})dW_u,$$

for $s\in[t,T]$. Therefore, for $T-t=o(1)$, we have $X_T^{t,x;\hat{q}} \approx Y_T^{t,y}$, where $Y^{t,y}$ solves the *uncontrolled* stochastic differential equation

$$Y_s^{t,y} = y + \int_t^s \sigma(u, Y_u^{t,y})dW_u,$$

for $s\in[t,T]$. Note that, since y is arbitrary, we readily obtain an upper bound of the solution $u(t,x)$ of (3.4), namely,

$$u(t,x) \le \inf_{y\in\mathbb{R}^d}\left\{(T-t)L(\frac{x-y}{T-t}) + \mathbb{E}[U(Y_T^{t,y})|\mathcal{F}_t]\right\}. \tag{3.5}$$

Furthermore, the convexity of H yields that L is also convex and, therefore, for any control process $q\in\mathbb{H}^2[t,T]$, we deduce that

$$\begin{aligned}
&\mathbb{E}\left[\int_t^T L(q_s)ds + U(X_T^{t,x;q})\middle|\, \mathcal{F}_t\right] \\
&\ge (T-t)L\left(\mathbb{E}\left[\frac{1}{T-t}\int_t^T q_u du\middle|\, \mathcal{F}_t\right]\right) + \mathbb{E}[U(X_T^{t,x;q})|\mathcal{F}_t] \\
&= (T-t)L\left(\mathbb{E}\left[\frac{x - X_T^{t,x;q} + \int_t^T \sigma(u, X_u^{t,x;q})dW_u}{T-t}\middle|\, \mathcal{F}_t\right]\right) + \mathbb{E}[U(X_T^{t,x;q})|\mathcal{F}_t] \\
&= (T-t)L\left(\frac{x - \mathbb{E}[X_T^{t,x;q}|\mathcal{F}_t]}{T-t}\right) + \mathbb{E}[U(X_T^{t,x;q})|\mathcal{F}_t].
\end{aligned}$$

Therefore, for $T-t=o(1)$, we have $X_T^{t,x;q} \approx Y_T^{t,\hat{y}}$, with $\hat{y} := \mathbb{E}[X_T^{t,x;q}|\mathcal{F}_t]$. Thus, we also obtain a lower bound of the solution $u(t,x)$ of (3.4), namely,

$$u(t,x) \geq \inf_{\hat{y}\in\mathbb{R}^d} \left\{ (T-t)L(\frac{x-\hat{y}}{T-t}) + \mathbb{E}[U(Y_T^{t,\hat{y}})|\mathcal{F}_t] \right\}. \tag{3.6}$$

Note that when σ degenerates to 0, inequalities (3.5) and (3.6) give us an equality, which is precisely the Hopf-Lax formula (3.3).

We now see how the above ideas can be combined to develop an approximation scheme for the original problem (3.1). Equation (3.1) can be "split" into a first-order nonlinear equation of Hamilton-Jacobi type and a linear parabolic equation. The solution of the former is represented via the Hopf-Lax formula and corresponds to the value function of a deterministic control problem. The solution of the latter corresponds to a conditional expectation of an uncontrolled diffusion and is given by the Feynman-Kac formula. The scheme is then naturally based on a backwards-in-time recursive combination of the Hopf-Lax and the Feynman-Kac formula; see (3.8) and (3.18) for further details.

We then establish the convergence of the scheme to the unique (viscosity) solution of (3.1) and determine the rate of convergence. We do this by establishing upper and lower bounds on the approximation error (Theorems 3.3.5 and 3.3.8, respectively). The main tools come from the *shaking coefficients technique* introduced by Krylov [43, 44] and the *optimal switching approximation* introduced by Barles and Jacobsen [3, 4].

While various arguments follow from adaptations of these techniques, the main difficulty is to derive a consistency error estimate. This is one of the key steps herein and it is precisely where the convexity of the Hamiltonian with respect to the gradient is used in an essential way. Specifically, we obtain this estimate by applying convex duality and using the properties of the optimizers in the related minimization problems (Proposition 3.2.5 (vi)). Using this estimate and the comparison result for the approximation scheme (Proposition 3.2.8), we in turn derive an upper bound for the approximation error by perturbing the coefficients of the equation. The lower bound for the approximation error is obtained by another layer of approximation of the equation by using an auxiliary optimal switching system.

The splitting approximation approach (fractional step, prediction and correction, etc.) is dated back to Marchuk [49] in the late 1960s. Its application to nonlinear PDEs was firstly proposed by Lions and Mercier [47] and has been subsequently used by many others. For semilinear parabolic equations related to problems in mathematical finance, splitting methods have been applied by Tourin [63] (see also more references therein). More recently, Nadtochiy and Zariphopoulou [50] proposed a splitting algorithm to the marginal HJB equation arising in optimal investment problems in a stochastic factor model and general utility functions. Henderson and Liang [33] proposed a splitting approach for utility indifference pricing in a multi-dimensional non-traded assets model with intertemporal default risk, and established its convergence rate. Tan [62] proposed a splitting method for a class of fully nonlinear degenerate parabolic PDEs and applied it to Asian options and commodity trading.

Finally, we mention that most of the existing algorithms (see, among others, Howard's finite difference scheme [8]) provide approximations at certain time grids and use interpolation for any time point in between. In contrast, the approximation scheme we propose uses interpolation only for time points in the last time-interval and approximates the solution at all the remaining time points by a uniform algorithm (see (3.17)). This ensures that the approximation solution satisfies a scheme equation (see (3.18)) in the whole domain rather than just at certain time grids. Furthermore, the connection between BSDEs and the semi-linear equations (3.1) relies on the requirement that the Hamiltonian has the peculiar form $H(t,x,p) = g(t,x,\sigma^T(t,x)p)$ where g is the driver of the associated BSDE (see [40]). In this sense, the approximation scheme we propose is a more general way of solving (3.1) than the commonly used time discretization algorithms based on BSDE techniques (see [10] and [14]). Conversely, the proposed scheme also provides a new algorithm to numerically solve the associated BSDE.

This chapter is organized as follows. In section 3.2 we introduce the monotone approximation scheme. In section 3.3, we prove its convergence rate using the shaking coefficients technique and optimal switching approximation. We then give a numerical example in section 3.4. Some technical proofs are provided in the appendix.

3.2 The monotone scheme with Hopf-Lax formula and splitting

We throughout assume the following conditions for equation (3.1).

Assumption 3.2.1 *(i) The diffusion coefficient σ, the drift coefficient b, and the terminal datum U have norms $|\sigma|_1, |b|_1, |U|_1 \leq M$, for some $M > 0$.*

(ii) The Hamiltonian $H(t,x,p) \in \mathcal{C}(Q_T \times \mathbb{R}^d)$ is convex in p, and satisfies the coercivity condition

$$\lim_{|p|\to\infty} \frac{H(t,x,p)}{|p|} = \infty,$$

uniformly in $(t,x) \in Q_T$. Moreover, for every p, $[H(\cdot,\cdot,p)]_{\mathcal{C}^{1/2,1}} \leq M$, and there exist two locally bounded functions H^* and $H_* : \mathbb{R}^d \to \mathbb{R}$ such that

$$H_*(p) = \inf_{(t,x)\in Q_T} H(t,x,p) \quad \textit{and} \quad H^*(p) = \sup_{(t,x)\in Q_T} H(t,x,p).$$

Unless state otherwise, we will then throughout this chapter denote by $C := C(T,M)$ some constant that depends only on T and M. Under the above assumptions, we have the following existence, uniqueness and regularity results for equation (3.1). Their proofs are provided in Appendix A.

Proposition 3.2.2 *Suppose that Assumption 3.2.1 is satisfied. Then, there exists a unique viscosity solution $u \in \mathcal{C}_b^1(\bar{Q}_T)$ of equation (3.1), with $|u|_1 \leq C$.*

3.2.1 The backward operator $\mathbf{S}_t(\Delta)$

Using that $H(t,x,p)$ is convex in p, we define its Legendre (convex dual) transform $L : Q_T \times \mathbb{R}^d \to \mathbb{R}$ given by

$$L(t,x,q) := \sup_{p\in\mathbb{R}^d} \{p \cdot q - H(t,x,p)\}. \tag{3.7}$$

For any t and Δ with $0 \leq t < t+\Delta \leq T$, and any $\phi \in \mathcal{C}_b(\mathbb{R}^d)$, we introduce the *backward operator* $\mathbf{S}_t(\Delta) : \mathcal{C}_b(\mathbb{R}^d) \to \mathcal{C}_b(\mathbb{R}^d)$,

$$\begin{cases} \mathbf{S}_t(\Delta)\phi(x) = \min_{y\in\mathbb{R}^d} \left\{ \Delta L\left(t, x, \frac{x-y}{\Delta}\right) + \mathbb{E}[\phi(Y^{t,y}_{t+\Delta})|\mathcal{F}_t] \right\}, & x \in \mathbb{R}^d, \\ Y^{t,y}_s = y + b(t,y)(s-t) + \sigma(t,y)(W_s - W_t), & s \in [t, t+\Delta], \end{cases} \tag{3.8}$$

on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbf{P})$, where W is an d-dimensional Brownian motion with its augmented filtration $\{\mathcal{F}_t\}_{t\geq 0}$.

We start with some auxiliary properties of H and L.

Proposition 3.2.3 *Suppose that Assumption 3.2.1 (ii) is satisfied. Then, the following assertions hold:*

(i) H is locally Lipschitz in p uniformly in $(t,x) \in Q_T$. Moreover, H is the Legendre transform of L, i.e.

$$H(t,x,p) = \sup_{q\in\mathbb{R}^d} \{p \cdot q - L(t,x,q)\}, \quad \textit{for } (t,x) \in Q_T$$

. (ii) The functions

$$L_*(q) := \sup_{p\in\mathbb{R}^d} \{p\cdot q - H^*(p)\} \quad \textit{and} \quad L^*(q) := \sup_{p\in\mathbb{R}^d} \{p\cdot q - H_*(p)\}$$

are locally bounded and satisfy, for $(t,x) \in Q_T$, $L_*(q) \leq L(t,x,q) \leq L^*(q)$.

(iii) For $(t,x) \in Q_T$, $L(t,x,q)$ is convex in q with $[L(\cdot,\cdot,q)]_{\mathcal{C}^{1/2,1}} \leq M$. Furthermore, it satisfies the coercivity condition

$$\lim_{|q|\to\infty} \frac{L(t,x,q)}{|q|} = \infty,$$

uniformly in $(t,x) \in Q_T$.

(iv) For each $(t,x) \in Q_T$ and $p, q \in \mathbb{R}^d$, there exist $p^*, q^* \in \mathbb{R}^d$ such that

$$L(t,x,q) = q\cdot p^* - H(t,x,p^*) \quad \textit{and} \quad H(t,x,p) = p\cdot q^* - L(t,x,q^*).$$

Furthermore, $|p^*| \leq \xi(|q|)$ and $|q^*| \leq \xi(|p|)$, for some real-valued increasing function $\xi(\cdot)$ independent of (t,x).

Proof. Parts (i) and (ii) are immediate and, thus, we only prove (iii) and (iv):

(iii) For fixed $(t,x) \in Q_T$, $q_1, q_2 \in \mathbb{R}^d$ and $\lambda \in [0,1]$, we have

$$\begin{aligned} L(t,x,\lambda q_1 + (1-\lambda)q_2) &= \sup_{p\in\mathbb{R}^d} \{(\lambda q_1 + (1-\lambda)q_2)\cdot p - H(t,x,p)\} \\ &\leq \lambda \sup_{p\in\mathbb{R}^d} \{q_1 \cdot p - H(t,x,p)\} + (1-\lambda) \sup_{p\in\mathbb{R}^d} \{q_2\cdot p - H(t,x,p)\} \\ &= \lambda L(t,x,q_1) + (1-\lambda)L(t,x,q_2). \end{aligned}$$

From the definition of L, we further have, for any $q \in \mathbb{R}^d$ and $(t,x) \neq (t',x') \in$

Q_T,

$$\frac{|L(t,x,q)-L(t',x',q)|}{|t-t'|^{1/2}+|x-x'|} \leq \sup_{p\in\mathbb{R}^d} \frac{|H(t,x,p)-H(t',x',p)|}{|t-t'|^{1/2}+|x-x'|} \leq M,$$

and thus $[L(\cdot,\cdot,q)]_{\mathcal{C}^{1/2,1}} \leq M$.

Next, for any $K>0$, we deduce, by setting $p = K\frac{q}{|q|}$, that

$$L(t,x,q) \geq q\cdot K\frac{q}{|q|} - H(t,x,K\frac{q}{|q|}) \geq K|q| - \sup_{r\in B(0,K)} H^*(r).$$

Dividing both sides by $|q|$ and sending $|q|\to\infty$, the coercivity condition for L follows.

(iv) From (i) and (ii), we deduce that L and H are symmetric to each other and, thus, we only establish the assertions for L. To this end, for each $(t,x)\in Q_T$, we obtain, by setting $p=0$ in (3.7), that $L(t,x,q)\geq -H(t,x,0)$. Therefore, it suffices to find a real-valued increasing function, say $\xi(\cdot)$, such that, if $|p|>\xi(|q|)$, then

$$p\cdot q - H(t,x,p) < -H(t,x,0).$$

Indeed, it follows from Assumption 3.2.1 (ii) that there exists a real-valued increasing function, say $K_H(y)$, such that, for any $(t,x)\in Q_T$ and $|p|\geq K_H(y)$, we have $\frac{H(t,x,p)}{|p|}\geq y$. Setting $\xi(x) := \max\{K_H(|H^*(0)|+x),1\}$, we deduce that, for $|p|>\xi(|q|)$,

$$p\cdot q - H(t,x,p) \leq |p|(|q| - \frac{H(t,x,p)}{|p|}) < |q| - (|H^*(0)|+|q|) \leq -H(t,x,0),$$

and we easily conclude. ■

Next, we show that the minimum in (3.8) is actually achieved, i.e. for any $\phi\in\mathcal{C}_b(\mathbb{R}^d)$, there always exists an associated minimizer y^*.

Proposition 3.2.4 *Suppose that Assumption 3.2.1 is satisfied. Then, for each t and Δ with $0\leq t<t+\Delta\leq T$, $x\in\mathbb{R}^d$ and $\phi\in\mathcal{C}_b(\mathbb{R}^d)$, there exists a minimizer $y^*\in\mathbb{R}^d$ such that*

$$\mathbf{S}_t(\Delta)\phi(x) = \Delta L\left(t,x,\frac{x-y^*}{\Delta}\right) + \mathbb{E}[\phi(Y^{t,y^*}_{t+\Delta})|\mathcal{F}_t].$$

Moreover,

$$\left|\frac{x-y^*}{\Delta}\right| \leq \xi(C[\phi]_{\mathcal{C}^1}), \tag{3.9}$$

for some real-valued increasing function $\xi(\cdot)$ independent of (t,x).

Proof. Let $q=\frac{x-y}{\Delta}$. Then $|q|\to\infty$ as $|y|\to\infty$. In turn, from Proposition 3.2.3 (iii), we deduce that, as $|y|\to\infty$,

$$\Delta L\left(t,x,\frac{x-y}{\Delta}\right) + \mathbb{E}[\phi(Y^{t,y}_{t+\Delta})|\mathcal{F}_t] = |x-y|\frac{L(t,x,q)}{|q|} + \mathbb{E}[\phi(Y^{t,y}_{t+\Delta})|\mathcal{F}_t] \to \infty.$$

Furthermore, using that the mapping $y \mapsto \Delta L(t, x, \frac{x-y}{\Delta}) + \mathbb{E}[\phi(Y^{t,y}_{t+\Delta})|\mathcal{F}_t]$ is continuous, we deduce that it must admit a minimizer $y^* \in \mathbb{R}^d$.

Next, we prove inequality (3.9). For $\phi \in \mathcal{C}^1_b(\mathbb{R}^d)$, following the same reasoning as in the proof of Proposition 3.2.3 (iv), it suffices to find a real-valued increasing function $\xi(\cdot)$ such that

$$\Delta L\left(t, x, q\right) + \mathbb{E}[\phi(Y^{t,x-\Delta q}_{t+\Delta})|\mathcal{F}_t] > \Delta L\left(t, x, 0\right) + \mathbb{E}[\phi(Y^{t,x}_{t+\Delta})|\mathcal{F}_t], \tag{3.10}$$

if $|q| > \xi(C[\phi]_{\mathcal{C}^1})$, for some constant $C > 0$ depending only on M and T. To prove this, note that Assumption 3.2.1 (i) on σ and b implies that

$$\begin{aligned}\mathbb{E}[\phi(Y^{t,x}_{t+\Delta})|\mathcal{F}_t] - \mathbb{E}[\phi(Y^{t,x-\Delta q}_{t+\Delta})|\mathcal{F}_t] &\le [\phi]_{\mathcal{C}^1}\mathbb{E}\left[\left|Y^{t,x}_{t+\Delta} - Y^{t,x-\Delta q}_{t+\Delta}\right| |\mathcal{F}_t\right] \\ &\le C[\phi]_{\mathcal{C}^1}\Delta|q|. \end{aligned} \tag{3.11}$$

On the other hand, from Proposition 3.2.3 (iv), there exists a real-valued increasing function, say $K_L(y)$, such that, for any $(t, x) \in Q_T$ and $|q| \ge K_L(y)$, we have $\frac{L(t,x,q)}{|q|} \ge y$. Setting $\xi(x) := \max\{K_L(|L^*(0)| + x), 1\}$, we then deduce that, for $|q| > \xi(C[\phi]_{\mathcal{C}^1})$,

$$\frac{L(t, x, q)}{|q|} > |L^*(0)| + C[\phi]_{\mathcal{C}^1} \ge \frac{L^*(0)}{|q|} + C[\phi]_{\mathcal{C}^1} \ge \frac{L(t, x, 0)}{|q|} + C[\phi]_{\mathcal{C}^1}.$$

Using the above inequality, together with (3.11), we obtain (3.10). Finally, the case $[\phi]_{\mathcal{C}^1} = \infty$ follows trivially. ■

Next, we derive some key properties of the backward operator $\mathbf{S}_t(\Delta)$.

Proposition 3.2.5 *Suppose that Assumption 3.2.1 is satisfied. Then, for each t and Δ with $0 \le t < t + \Delta \le T$, the operator $\mathbf{S}_t(\Delta)$ has the following properties:*

(i) (Constant preserving) For any $\phi \in \mathcal{C}_b(\mathbb{R}^d)$ and $c \in \mathbb{R}$,

$$\mathbf{S}_t(\Delta)(\phi + c) = \mathbf{S}_t(\Delta)\phi + c.$$

(ii) (Monotonicity) For any $\phi, \psi \in \mathcal{C}_b(\mathbb{R}^d)$,

$$\mathbf{S}_t(\Delta)\phi - \mathbf{S}_t(\Delta)\psi \le \sup_{x \in \mathbb{R}^d} (\phi - \psi)(x).$$

(iii) (Concavity) For any $\phi \in \mathcal{C}_b(\mathbb{R}^d)$, $\mathbf{S}_t(\Delta)\phi$ is concave in ϕ.

(iv) (Stability) For any $\phi \in \mathcal{C}_b(\mathbb{R}^d)$,

$$|\mathbf{S}_t(\Delta)\phi|_0 \le C\Delta + |\phi|_0,$$

where $C = \max\{|L^*(0)|, |H^*(0)|\}$, with L^* and H^* as in Proposition 3.2.3 (ii) and Assumption 3.2.1 (ii). Therefore, the operator $\mathbf{S}_t(\Delta)$ is indeed a mapping from $\mathcal{C}_b(\mathbb{R}^d)$ to $\mathcal{C}_b(\mathbb{R}^d)$.

(v) For any $\phi \in \mathcal{C}^1_b(\mathbb{R}^d)$, there exists a constant C depending only on $[\phi]_{\mathcal{C}^1}$, M and T, such that

$$|\mathbf{S}_t(\Delta)\phi - \phi|_0 \le C\sqrt{\Delta}.$$

(vi) For any $\phi \in \mathcal{C}_b^\infty(\mathbb{R}^d)$, define

$$\mathcal{E}(t,\Delta,\phi) := \left| \frac{\phi - \mathbf{S}_t(\Delta)\phi}{\Delta} - \mathbf{L}_t\phi \right|_0 , \tag{3.12}$$

where the operator $\mathbf{L}_t$ is given by

$$\mathbf{L}_t\phi(x) = -\frac{1}{2} tr\left(\sigma\sigma^T(t,x) D_x^2\phi(x)\right) - b(t,x)\cdot D_x\phi(x) + H(t,x,D_x\phi(x)).$$

Then,

$$\mathcal{E}(t,\Delta,\phi) \leq \; C\Delta\left(|D_x^4\phi|_0 + \mathcal{R}(\phi)\right),$$

where the constant C depends only on $[\phi]_{\mathcal{C}^1}$, M and T, and $\mathcal{R}(\phi)$ represents the "insignificant" terms containing the derivatives of ϕ up to third order.

Proof. Parts (i)-(iii) are immediate. We only prove (iv)-(vi) and, in particular, for the case $d = 1$, since the general case follows along similar albeit more complicated arguments.

(iv) Choosing $y = x$ in (3.8) gives

$$\mathbf{S}_t(\Delta)\phi(x) \leq \Delta L^*(0) + |\phi|_0. \tag{3.13}$$

It follows from the definition of L_* in Proposition 3.2.3 (ii) that $L_*(q) \geq -H^*(0) \geq -|H^*(0)|$, for $q \in \mathbb{R}^d$. In turn, Proposition 3.2.4 further yields

$$\begin{aligned}\mathbf{S}_t(\Delta)\phi(x) &= \Delta L(t,x,\frac{x-y^*}{\Delta}) + \mathbb{E}[\phi(Y_{t+\Delta}^{t,y^*})|\mathcal{F}_t] \\ &\geq \Delta L_*(\frac{x-y^*}{\Delta}) - |\phi|_0 \\ &\geq \; -\Delta|H^*(0)| - |\phi|_0. \end{aligned}\tag{3.14}$$

The assertion then follows by combining (3.13) and (3.14).

(v) From Proposition 3.2.3 (ii) and Proposition 3.2.4, we deduce that

$$\begin{aligned}|\mathbf{S}_t(\Delta)\phi(x) - \phi(x)| = &\left|\Delta L(t,x,\frac{x-y^*}{\Delta}) + \mathbb{E}[\phi(Y_{t+\Delta}^{t,y^*}) - \phi(x)|\mathcal{F}_t]\right| \\ &\leq \Delta\max\left\{|L^*(\frac{x-y^*}{\Delta})|, |L_*(\frac{x-y^*}{\Delta})|\right\} + [\phi]_{\mathcal{C}^1}\mathbb{E}\left[\left|Y_{t+\Delta}^{t,y^*} - x\right| |\mathcal{F}_t\right] \\ &\leq C\Delta + (C\Delta + M\Delta + CM\sqrt{\Delta})[\phi]_{\mathcal{C}^1} \leq C\sqrt{\Delta},\end{aligned}$$

where the constant C depends only on $[\phi]_{\mathcal{C}^1}$, M and T.

(vi) For $(t,x) \in [0, T-\Delta]\times\mathbb{R}$, let $q^* \in \mathbb{R}$ be such that

$$H(t,x,\partial_x\phi(x)) = \max_{q\in\mathbb{R}}\{q\partial_x\phi(x) - L(t,x,q)\} = q^*\partial_x\phi(x) - L(t,x,q^*).$$

From Proposition 3.2.3 (iv), we have $|q^*| \leq \xi(|\partial_x\phi(x)|) \leq C$, where the constant C depends only on $[\phi]_{\mathcal{C}^1}$, M and T.

Choosing $y = x - \Delta q^*$ in (3.8) and applying Itô's formula to $\phi(Y_{t+\Delta}^{t,x-\Delta q^*})$

yield

$$
\begin{aligned}
&\phi(x) - \mathbf{S}_t(\Delta)\phi(x) - \Delta\mathbf{L}_t\phi(x) \\
\geq\ &\phi(x) - \Delta L(t,x,q^*) - \phi(x-\Delta q^*) - \mathbb{E}[\phi(Y_{t+\Delta}^{t,x-\Delta q^*}) - \phi(x-\Delta q^*)|\mathcal{F}_t] - \Delta\mathbf{L}_t\phi(x) \\
=\ &(\phi(x) - \phi(x-\Delta q^*) - \Delta q^*\partial_x\phi(x)) \\
&- \left(\mathbb{E}\left[\int_t^{t+\Delta}\left(b(t,y)\partial_x\phi(Y_s^{t,x-\Delta q^*}) + \frac{1}{2}|\sigma(t,y)|^2\partial_{xx}\phi(Y_s^{t,x-\Delta q^*})\right)ds|\mathcal{F}_t\right]\right. \\
&\left.- \Delta b(t,x)\partial_x\phi(x) - \frac{1}{2}\Delta|\sigma(t,x)|^2\partial_{xx}\phi(x)\right) := (I) - (II).
\end{aligned}
$$

Next, we obtain a lower and an upper bound for terms (I) and (II), respectively. To this end, Taylor's expansion (2.23) yields

$$
\begin{aligned}
&\phi(x) - \phi(x-\Delta q^*) - \Delta q^*\partial_x\phi(x) \\
\geq\ &-(\Delta q^*)^2|\partial_{xx}\phi|_0 \geq -C\Delta^2|\partial_{xx}\phi|_0.
\end{aligned}
\tag{3.15}
$$

For term (II), applying Itô's formula to $\partial_x\phi(Y_s^{t,x-\Delta q^*})$ and $\partial_{xx}\phi(Y_s^{t,x-\Delta q^*})$ gives

$$
\begin{aligned}
&\mathbb{E}\left[\partial_x\phi(Y_s^{t,x-\Delta q^*})|\mathcal{F}_t\right] \\
=\ &\partial_x\phi(y) + \int_t^s \mathbb{E}\left[b(t,y)\partial_{xx}\phi(Y_u^{t,x-\Delta q^*}) + \frac{1}{2}|\sigma(t,y)|^2\partial_{xxx}\phi(Y_u^{t,x-\Delta q^*})|\mathcal{F}_t\right]du,
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{E}\left[\partial_{xx}\phi(Y_s^{t,x-\Delta q^*})|\mathcal{F}_t\right] \\
=\ &\partial_{xx}\phi(y) + \int_t^s \mathbb{E}\left[b(t,y)\partial_{xxx}\phi(Y_u^{t,x-\Delta q^*}) + \frac{1}{2}|\sigma(t,y)|^2\partial_{xxxx}\phi(Y_u^{t,x-\Delta q^*})|\mathcal{F}_t\right]du.
\end{aligned}
$$

Keeping the terms involving the derivatives of ϕ and using Assumption 3.2.1 on b and σ, we further have

$$
\begin{aligned}
&\mathbb{E}\left[\int_t^{t+\Delta}\left(b(t,y)\partial_x\phi(Y_s^{t,x-\Delta q^*}) + \frac{1}{2}|\sigma(t,y)|^2\partial_{xx}\phi(Y_s^{t,x-\Delta q^*})\right)ds|\mathcal{F}_t\right] \\
&- \Delta b(t,x)\partial_x\phi(x) - \frac{1}{2}\Delta|\sigma(t,x)|^2\partial_{xx}\phi(x) \\
\leq\ &C\Delta^2(|\partial_x\phi|_0 + |\partial_{xx}\phi|_0 + |\partial_{xxx}\phi|_0 + |\partial_{xxxx}\phi|_0).
\end{aligned}
\tag{3.16}
$$

In turn, combining estimates (3.15) and (3.16) above, we deduce that

$$
\frac{\phi(x) - \mathbf{S}_t(\Delta)\phi(x)}{\Delta} - \mathbf{L}_t\phi(x) \geq -C\Delta(|\partial_x\phi|_0 + |\partial_{xx}\phi|_0 + |\partial_{xxx}\phi|_0 + |\partial_{xxxx}\phi|_0),
$$

where the constant C depends only on $[\phi]_{\mathcal{C}^1}$, M and T.

To prove the reverse inequality, we work as follows. For $(t,x) \in [0,T-\Delta]\times\mathbb{R}$, let $y^* \in \mathbb{R}$ be the minimizer in (3.8) and set $p^* := \frac{x-y^*}{\Delta}$. Then, it follows from Proposition 3.2.4 that $|p^*| \leq C$, where the constant C depends only on $[\phi]_{\mathcal{C}^1}$,

M and T. Similar calculations as above then yield

$$
\begin{aligned}
&\phi(x) - \mathbf{S}_t(\Delta)\phi(x) - \Delta \mathbf{L}_t \phi(x) \\
&= \phi(x) - \Delta L(t,x,p^*) - \phi(x-\Delta p^*) - \mathbb{E}[\phi(Y_{t+\Delta}^{t,x-\Delta p^*}) - \phi(x-\Delta p^*)|\mathcal{F}_t] - \Delta \mathbf{L}_t \phi(x) \\
&= \Delta \left(p^* \partial_x \phi(x) - L(t,x,p^*)\right) - \Delta H(t,x,\partial_x\phi(x)) - \int_x^{x-\Delta p^*} \left(\int_x^s \partial_{xx}\phi(u)du\right) ds \\
&\quad - \left(\mathbb{E}\left[\int_t^{t+\Delta} \left(b(t,y)\partial_x\phi(Y_s^{t,x-\Delta p^*}) + \frac{1}{2}|\sigma(t,y)|^2 \partial_{xx}\phi(Y_s^{t,x-\Delta p^*})\right) ds|\mathcal{F}_t\right]\right. \\
&\quad \left. - \Delta b(t,x)\partial_x\phi(x) - \frac{1}{2}\Delta|\sigma(t,x)|^2\partial_{xx}\phi(x)\right) \\
&\leq C\Delta^2(|\partial_x\phi|_0 + |\partial_{xx}\phi|_0 + |\partial_{xxx}\phi|_0 + |\partial_{xxxx}\phi|_0),
\end{aligned}
$$

for some constant C depending only on $[\phi]_{\mathcal{C}^1}$, M and T. We easily conclude. ■

3.2.2 The approximation scheme

We now introduce the approximation scheme for equation (3.1). For $\Delta \in (0,T)$ and $(t,x) \in \bar{Q}_{T-\Delta}$, we introduce the iterative algorithm

$$u^\Delta(t,x) = \mathbf{S}_t(\Delta)u^\Delta(t+\Delta,\cdot)(x), \tag{3.17}$$

with $u^\Delta(T,\cdot) = U(\cdot)$ and $\mathbf{S}_t(\Delta)$ defined in (3.8). The values between $T-\Delta$ and T are obtained by a standard linear interpolation.

Specifically, the approximation scheme is given by

$$\begin{cases} S(\Delta,t,x,u^\Delta(t,x),u^\Delta(t+\Delta,\cdot)) = 0 & \text{in } \bar{Q}_{T-\Delta}; \\ u^\Delta(t,x) = g^\Delta(t,x) & \text{in } \bar{Q}_T \backslash \bar{Q}_{T-\Delta}, \end{cases} \tag{3.18}$$

where $S : (0,T) \times \bar{Q}_{T-\Delta} \times \mathbb{R} \times \mathcal{C}_b(\mathbb{R}^d) \to \mathbb{R}$ and $g^\Delta : \bar{Q}_T \backslash \bar{Q}_{T-\Delta} \to \mathbb{R}$ are defined, respectively, by

$$S(\Delta,t,x,p,v) = \frac{p - \mathbf{S}_t(\Delta)v(x)}{\Delta} \tag{3.19}$$

and

$$g^\Delta(t,x) = \omega_1(t)U(x) + \omega_2(t)\mathbf{S}_{T-\Delta}(\Delta)U(x), \tag{3.20}$$

with $\omega_1(t) = (t+\Delta-T)/\Delta$ and $\omega_2(t) = (T-t)/\Delta$ being the linear interpolation weights.

Note that when $T-\Delta < t \leq T$, the approximation term g^Δ corresponds to the usual linear interpolation between $T-\Delta$ and T. When $t = T-\Delta$, we have $\omega_1(t) = 0$ and $\omega_2(t) = 1$ and, thus, $g^\Delta(T-\Delta,x) = u^\Delta(T-\Delta,x)$.

We first prove the stability property of the approximation scheme (3.18).

Lemma 3.2.6 *(Stability) Suppose that Assumption 2.1 is satisfied. Then, the approximation scheme (3.18) admits a unique solution $u^\Delta \in \mathcal{C}_b(\bar{Q}_T)$, with $|u^\Delta|_0 \leq C$.*

Proof. By the stability property (iv) in Proposition 3.2.5, we have that $\mathbf{S}_t(\Delta)\phi$ is uniformly bounded if so is ϕ. Therefore, (3.17) is always well defined in $\bar{Q}_{T-\Delta}$, which yields the existence and uniqueness of the solution u^Δ. Furthermore, for $0 \le t \le T-\Delta$, $|u^\Delta(t,\cdot)|_0 \le C\Delta + |u^\Delta(t+\Delta,\cdot)|_0$. By backward induction and the definition of g^Δ in (3.18), we conclude that

$$|u^\Delta|_0 \le CT + \sup_{t\in(T-\Delta,T]} |g^\Delta(t,\cdot)|_0 \le C.$$

■

Using the properties of $\mathbf{S}_t(\Delta)$ established in Proposition 3.2.5, we next obtain the following key properties of the approximation scheme (3.18).

Proposition 3.2.7 *Suppose that Assumption 3.2.1 is satisfied. Then, for each t and Δ with $0 \le t < t+\Delta \le T$, $x \in \mathbb{R}^d$, $p \in \mathbb{R}$ and $v \in \mathcal{C}_b(\mathbb{R}^d)$, the approximation scheme $S(\Delta, t, x, p, v)$ has the following properties:*

(i) (Monotonicity) For any $c_1, c_2 \in \mathbb{R}$ and $u \in \mathcal{C}_b(\mathbb{R}^d)$ with $u \le v$,

$$S(\Delta, t, x, p + c_1, u + c_2) \ge S(\Delta, t, x, p, v) + \frac{c_1 - c_2}{\Delta}.$$

(ii) (Convexity) $S(\Delta, t, x, p, v)$ is convex in p and v.

(iii) (Consistency) For any $\phi \in \mathcal{C}_b^\infty(\bar{Q}_T)$, there exists a constant C, depending only on $[\phi]_{\mathcal{C}^{1/2,1}}$, M and T, such that

$$\begin{aligned}
&|-\partial_t\phi(t,x) + \mathbf{L}_t\phi(t,x) - S(\Delta, t, x, \phi(t,x), \phi(t+\Delta,\cdot))| \\
&\le C\Delta\left(|\partial_t^2\phi|_0 + |D_x^4\phi|_0 + |\partial_t D_x^2\phi|_0 + \mathcal{R}(\phi)\right).
\end{aligned} \tag{3.21}$$

Proof. Parts (i) and (ii) follow easily from Proposition 3.2.5. Similarly to the proof of Proposition 3.2.5(vi), we only prove (iii) for the case $d = 1$. To this end, we split the consistency error into three parts. Specifically,

$$\begin{aligned}
&|-\partial_t\phi(t,x) + \mathbf{L}_t\phi(t,x) - S(\Delta, t, x, \phi(t,x), \phi(t+\Delta,\cdot))| \\
&\le \mathcal{E}(t, \Delta, \phi(t+\Delta,\cdot)) + |\phi(t+\Delta,x) - \phi(t,x) - \Delta\partial_t\phi(t,x)|\Delta^{-1} \\
&\quad + |\mathbf{L}_t\phi(t,x) - \mathbf{L}_t\phi(t+\Delta,x)| := (I) + (II) + (III),
\end{aligned}$$

where $\mathcal{E}$ was defined in (3.12). For term (I), Proposition 3.2.5 (vi) yields

$$\begin{aligned}
\mathcal{E}(t, \Delta, \phi(t+\Delta,\cdot)) &\le C\Delta\left(|\partial_{xxxx}\phi(t+\Delta,\cdot)|_0 + \mathcal{R}(\phi(t+\Delta,\cdot))\right) \\
&\le C\Delta\left(|\partial_{xxxx}\phi|_0 + \mathcal{R}(\phi)\right),
\end{aligned} \tag{3.22}$$

for some constant C depending only on $[\phi]_{\mathcal{C}^{1/2,1}}$, M and T. For term (II), Taylor's expansion (2.23) gives

$$|\phi(t+\Delta,x) - \phi(t,x) - \Delta\partial_t\phi(t,x)|\Delta^{-1} \le \Delta|\partial_{tt}\phi|_0. \tag{3.23}$$

Finally, for term (III), we have from Assumption 3.2.1(i) and Proposition 3.2.3(i) that

$$\begin{aligned}
&|\mathbf{L}_t\phi(t,x) - \mathbf{L}_t\phi(t+\Delta,x)| \\
&\le C(|\partial_{xx}\phi(t,x) - \partial_{xx}\phi(t+\Delta,x)| + |\partial_x\phi(t,x) - \partial_x\phi(t+\Delta,x)|) \\
&\quad + |H(t,x,\partial_x\phi(t,x)) - H(t,x,\partial_x\phi(t+\Delta,x))| \\
&\le C\Delta(|\partial_{xxt}\phi|_0 + |\partial_{xt}\phi|_0),
\end{aligned} \tag{3.24}$$

for some constant C depending only on $[\phi]_{\mathcal{C}^{1/2,1}}$ and M. Combining estimates (3.22)-(3.24), we easily conclude. ■

The following comparison property for the approximation scheme (3.18) will be used frequently in the next section.

Proposition 3.2.8 *Suppose that Assumption 2.1 is satisfied, and that $u, v \in \mathcal{C}_b(\bar{Q}_T)$ are such that*

$$\begin{aligned} S(\Delta, t, x, u, u(t+\Delta, \cdot)) &\le h_1 \quad \text{in } \bar{Q}_{T-\Delta}; \\ S(\Delta, t, x, v, v(t+\Delta, \cdot)) &\ge h_2 \quad \text{in } \bar{Q}_{T-\Delta}, \end{aligned}$$

for some h_1, $h_2 \in \mathcal{C}_b(\bar{Q}_{T-\Delta})$. Then,

$$u - v \le \sup_{\bar{Q}_T \backslash \bar{Q}_{T-\Delta}} (u-v)^+ + (T-t) \sup_{\bar{Q}_{T-\Delta}} (h_1 - h_2)^+ \quad \text{in } \bar{Q}_T. \tag{3.25}$$

Proof. The inequality (3.25) holds obviously in $\bar{Q}_T \backslash \bar{Q}_{T-\Delta}$. In $\bar{Q}_{T-\Delta}$, we have

$$u(t,x) \le \mathbf{S}_t(\Delta) u(t+\Delta, \cdot)(x) + \Delta h_1;$$

$$v(t,x) \ge \mathbf{S}_t(\Delta) v(t+\Delta, \cdot)(x) + \Delta h_2.$$

Combining the above two inequalities and using Proposition 3.2.5(ii) yields

$$\begin{aligned} (u-v)(t,x) &\le \sup_{x \in \mathbb{R}^d} (u-v)(t+\Delta, x) + \Delta(h_1 - h_2) \\ &\le \sup_{x \in \mathbb{R}^d} (u-v)(t+\Delta, x) + \Delta \sup_{\bar{Q}_{T-\Delta}} (h_1 - h_2)^+ \end{aligned}$$

By induction, we have that there exists some $n \in \mathbb{N}$ such that $T - \Delta < t + n\Delta \le T$ and

$$\begin{aligned} (u-v)(t,x) &\le \sup_{x \in \mathbb{R}^d} (u-v)(t+n\Delta, x) + n\Delta \sup_{\bar{Q}_{T-\Delta}} (h_1 - h_2)^+ \\ &\le \sup_{\bar{Q}_T \backslash \bar{Q}_{T-\Delta}} (u-v)^+ + (T-t) \sup_{\bar{Q}_{T-\Delta}} (h_1 - h_2)^+ \end{aligned}$$

■

Corollary 3.2.9 *Consider only the grid $\bar{\mathcal{G}}_T^\Delta \subset \bar{Q}_T$, and let $\mathcal{G}_T^\Delta := \bar{\mathcal{G}}_T^\Delta \backslash \{t = T\}$ be the grid before terminal time T. Suppose that $u, v \in \mathcal{C}_b(\bar{Q}_T)$ are such that*

$$\begin{aligned} S(\Delta, t, x, u, u(t+\Delta, \cdot)) &\le h_1 \quad \text{in } \mathcal{G}_T^\Delta; \\ S(\Delta, t, x, v, v(t+\Delta, \cdot)) &\ge h_2 \quad \text{in } \mathcal{G}_T^\Delta, \end{aligned}$$

for some h_1, $h_2 \in \mathcal{C}_b(\mathcal{G}_T^\Delta)$. Then,

$$u - v \le |(u(T,\cdot) - v(T,\cdot))^+|_0 + (T-t)|(h_1 - h_2)^+|_0 \quad \text{in } \bar{\mathcal{G}}_T^\Delta. \tag{3.26}$$

3.3 Convergence rate of the approximation scheme

In this section, we establish the convergence rate of the approximate solution u^Δ to the viscosity solution u of equation (3.1). We start with the approximation error estimate $u - u^\Delta$ in the final time interval $\bar{Q}_T \backslash \bar{Q}_{T-\Delta}$, where the value of u^Δ involves only a one-step approximation and some linear interpolation. Therefore, the bound for the approximation error in this domain can be easily obtained by the properties of the backward operator $\mathbf{S}_t(\Delta)$ and the regularity results of u.

Lemma 3.3.1 *Suppose that Assumption 3.2.1 holds. Let $u^\Delta \in \mathcal{C}_b(\bar{Q}_T)$ satisfy the approximation scheme (3.18) and $u \in \mathcal{C}_b^1(\bar{Q}_T)$ be the unique viscosity solution of equation (3.1). Then,*

$$|u - u^\Delta| \leq C\sqrt{\Delta} \quad \text{in } \bar{Q}_T \backslash \bar{Q}_{T-\Delta}. \tag{3.27}$$

Proof. From (3.18), we have, for $(t, x) \in \bar{Q}_T \backslash \bar{Q}_{T-\Delta}$,

$$\begin{aligned} |u(t,x) - u^\Delta(t,x)| &= |u(t,x) - g^\Delta(t,x)| \\ &= |u(t,x) - u(T,x) + \omega_2(t)(U(x) - \mathbf{S}_{T-\Delta}(\Delta)U(x))| \\ &\leq |u(t,x) - u(T,x)| + |U(x) - \mathbf{S}_{T-\Delta}(\Delta)U(x)| \\ &\leq C(\sqrt{|T-t|} + \sqrt{\Delta}) \leq C\sqrt{\Delta}, \end{aligned}$$

where the second to last inequality follows from the regularity property of the solution u (cf. Proposition 3.2.2) and property (v) of the operator $\mathbf{S}_t(\Delta)$ (cf. Proposition 3.2.5). ■

Next, we derive the bound of approximation error within the whole domain $\bar{Q}_T$. We start with the special case when (3.1) has a unique smooth solution u with bounded derivatives of any order.

Theorem 3.3.2 *Suppose that Assumption 3.2.1 is satisfied and that equation (3.1) admits a unique smooth solution $u \in \mathcal{C}_b^\infty(\bar{Q}_T)$. Then,*

$$|u - u^\Delta| \leq C\Delta \quad \text{in } \bar{Q}_T.$$

Proof. Using that $u \in \mathcal{C}_b^\infty(\bar{Q}_T)$, the consistency error estimate (3.21) yields

$$\begin{aligned} &|S(\Delta, t, x, u(t,x), u(t+\Delta, \cdot))| \\ &\leq C\Delta \left(|\partial_t^2 u|_0 + |D_x^4 u|_0 + |\partial_t D_x^2 u|_0 + \mathcal{R}(u) \right) \leq C\Delta, \end{aligned}$$

for $(t, x) \in \bar{Q}_{T-\Delta}$. On the other hand, from the definition of the approximation scheme (3.18), we have

$$S(\Delta, t, x, u^\Delta(t,x), u^\Delta(t+\Delta, \cdot)) = 0,$$

for $(t, x) \in \bar{Q}_{T-\Delta}$. In turn, the comparison property in Proposition 3.2.8 yields

$$u(t,x) - u^\Delta(t,x) \leq \sup_{(t,x) \in \bar{Q}_T \backslash \bar{Q}_{T-\Delta}} (u(t,x) - u^\Delta(t,x))^+ + (T-t)C\Delta,$$

and

$$u^\Delta(t,x) - u(t,x) \leq \sup_{(t,x) \in \bar{Q}_T \backslash \bar{Q}_{T-\Delta}} (u^\Delta(t,x) - u(t,x))^+ + (T-t)C\Delta,$$

for $(t,x) \in \bar{Q}_T$. It is left to prove that $|u - u^\Delta| < C\Delta$ in $\bar{Q}_T \backslash \bar{Q}_{T-\Delta}$. Indeed, the corollary comparison result in Corollary 3.2.9 gives

$$|u(T-\Delta, x) - u^\Delta(T-\Delta, x)| \leq (T - (T-\Delta))C\Delta = C\Delta^2,$$

and thus in $\bar{Q}_T \backslash \bar{Q}_{T-\Delta}$,

$$\begin{aligned}
&|u(t,x) - u^\Delta(t,x)| = |u(t,x) - \omega_1(t)U(x) + \omega_2(t)u^\Delta(T-\Delta, x)| \\
&\leq \omega_1(t)|u(t,x) - U(x)| + \omega_2(t)\left(|u(t,x) - u(T-\Delta,x)| + |u(T-\Delta,x) - u^\Delta(T-\Delta,x)|\right) \\
&\leq \omega_1(t)[u]_{\mathcal{C}^{1,1}}\Delta + \omega_2(t)[u]_{\mathcal{C}^{1,1}}\Delta + \omega_2(t)C\Delta^2 \\
&\leq C\Delta
\end{aligned}$$

We easily conclude. ■

In general, the above result might not hold as (3.1) only admits a viscosity solution $u \in \mathcal{C}_b^1(\bar{Q}_T)$ due to possible degeneracies. We follow a similar idea as in Section 2.4.2 to approximate the viscosity solution u by a sequence of smooth sub- and supersolutions u_ε and, in turn, compare them with u^Δ using the comparison property of the approximation scheme developed in Proposition 3.2.8. We carry out this procedure next.

3.3.1 Upper bound for the approximation error

We first derive an upper bound for the approximation error $u - u^\Delta$. We do so by first constructing a sequence of smooth subsolutions to equation (3.1) by perturbing its coefficients. As we mentioned in Section 1.1, this approach, known as the *shaking coefficients technique*, was initially proposed by Krylov [43, 44], and further developed by Barles and Jakobsen [2, 39].

To this end, for small enough $\varepsilon \geq 0$, we extend the functions $f := \sigma, b$ and H to $Q_{T+\varepsilon^2}^{-\varepsilon^2} := [-\varepsilon^2, T+\varepsilon^2) \times \mathbb{R}^d$ and $Q_{T+\varepsilon^2}^{-\varepsilon^2} \times \mathbb{R}^d$, respectively, so that Assumption 3.2.1 still holds. We then define $f^\theta(t,x) := f(t+\tau, x+e)$ and $H^\theta(t,x,p) := H(t+\tau, x+e, p)$, where $\theta = (\tau, e)$ with $\theta \in \Theta^\varepsilon := [-\varepsilon^2, 0] \times \varepsilon B(0,1)$. We then consider the perturbed version of equation (3.1), namely,

$$\begin{cases} -\partial_t u^\varepsilon + \sup\limits_{\theta \in \Theta^\varepsilon} g^\theta(t,x,D_x u^\varepsilon, D_x^2 u^\varepsilon) = 0 & \text{in } Q_{T+\varepsilon^2}; \\ u^\varepsilon(T+\varepsilon^2, x) = U(x) & \text{in } \mathbb{R}^d. \end{cases} \tag{3.28}$$

with

$$g^\theta(t,x,p,X) = -\frac{1}{2}\mathrm{tr}\left(\sigma^\theta {\sigma^\theta}^T(t,x)X\right) - b^\theta(t,x) \cdot p + H^\theta(t,x,p).$$

Note that when the perturbation parameter $\varepsilon = 0$, equations (3.28) and (3.1) coincide.

We establish existence, uniqueness and regularity results for the HJB equation (3.28), and a comparison between u and u^ε. Their proofs are provided in Appendix A.

Proposition 3.3.3 *Suppose that Assumption 3.2.1 is satisfied. Then, there exists a unique viscosity solution $u^\varepsilon \in \mathcal{C}_b^1(\bar{Q}_{T+\varepsilon^2})$ of equation (3.28), with*

$|u^\varepsilon|_1 \le C$. *Moreover,*

$$|u - u^\varepsilon| \le C\varepsilon \quad \text{in } \bar{Q}_T. \tag{3.29}$$

Next, we regularize u^ε by the same mollification procedure as introduced in Section 2.4.1. For this, let $\rho(t,x)$ be a $\mathbb{R}_+$-valued smooth function with compact support $\{-1 < t < 0\} \times \{|x| < 1\}$ and mass 1, and introduce the sequence of mollifiers ρ_ε,

$$\rho_\varepsilon(t,x) := \frac{1}{\varepsilon^{d+2}}\rho\left(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\right). \tag{3.30}$$

For $(t,x) \in \bar{Q}_T$, we then define

$$u_\varepsilon(t,x) = u^\varepsilon * \rho_\varepsilon(t,x) = \int_{-\varepsilon^2<\tau<0}\int_{|e|<\varepsilon} u^\varepsilon(t-\tau, x-e)\rho_\varepsilon(\tau,e)ded\tau.$$

Similar to Section 2.4.1, we have $u_\varepsilon \in \mathcal{C}_b^\infty(\bar{Q}_T)$,

$$|u^\varepsilon - u_\varepsilon|_0 \le C\varepsilon, \tag{3.31}$$

and, for positive integers i and multi-index j,

$$|\partial_t^i D_x^j u_\varepsilon|_0 \le C\varepsilon^{1-2i-|j|}, \tag{3.32}$$

where the constant C is independent of ε.

Lemma 3.3.4 *u_ε is a (classical) subsolution of (3.1) in Q_T.*

Proof. We observe that the function $u^\varepsilon(t-\tau, x-e)$, $(t,x) \in Q_T$, is a viscosity subsolution of equation (3.1) in Q_T, for any $(\tau, e) \in \Theta^\varepsilon$. On the other hand, a Riemann sum approximation shows that $u_\varepsilon(t,x)$ can be viewed as the limit of convex combinations of $u^\varepsilon(t-\tau,x-e)$, for $(\tau,e) \in \Theta^\varepsilon$. Since the equation in (3.1) is convex in $D_x u$, and linear in $\partial_t u$ and $D_x^2 u$, the convex combinations of $u^\varepsilon(t-\tau,x-e)$ are also subsolutions of (3.1) in Q_T. Using the stability of viscosity solutions (see Proposition 2.2.6), we deduce that $u_\varepsilon(t,x)$ is also a subsolution of (3.1) in Q_T. ■

We are now ready to establish an upper bound for the approximation error.

Theorem 3.3.5 *Suppose that Assumption 3.2.1 holds. Let $u^\Delta \in \mathcal{C}_b(\bar{Q}_T)$ satisfy the approximation scheme (3.18) and $u \in \mathcal{C}_b^1(\bar{Q}_T)$ be the unique viscosity solution of equation (3.1). Then,*

$$u - u^\Delta \le C\Delta^{\frac{1}{4}} \quad \text{in } \bar{Q}_T.$$

Proof. Substituting u_ε into the consistency error estimate (3.21) and using (3.32) give

$$\begin{aligned}&|-\partial_t u_\varepsilon(t,x) + \mathbf{L}_t u_\varepsilon(t,x) - S(\Delta,t,x,u_\varepsilon(t,x),u_\varepsilon(t+\Delta,\cdot))|\\ &\le C\Delta\left(|\partial_t^2 u_\varepsilon|_0 + |D_x^4 u_\varepsilon|_0 + |\partial_t D_x^2 u_\varepsilon|_0 + \mathcal{R}(u_\varepsilon)\right) \le C\Delta\varepsilon^{-3},\end{aligned}$$

for $(t,x) \in \bar{Q}_{T-\Delta}$. Since u_ε is a subsolution of (3.1) in Q_T, we have

$$S(\Delta,t,x,u_\varepsilon(t,x),u_\varepsilon(t+\Delta,\cdot)) \le C\Delta\varepsilon^{-3},$$

for $(t,x) \in \bar{Q}_{T-\Delta}$. Furthermore, by the definition of the approximation scheme (3.18), we also have

$$S(\Delta, t, x, u^{\Delta}(t,x), u^{\Delta}(t+\Delta, \cdot)) = 0,$$

for $(t,x) \in \bar{Q}_{T-\Delta}$. In turn, the comparison result in Proposition 3.2.8 implies

$$u_{\varepsilon} - u^{\Delta} \leq \sup_{\bar{Q}_T \backslash \bar{Q}_{T-\Delta}} (u_{\varepsilon} - u^{\Delta})^{+} + C(T-t)\Delta\varepsilon^{-3} \quad \text{in } \bar{Q}_T.$$

Next, using estimates (3.29) and (3.31), we obtain that $|u - u_{\varepsilon}| \leq C\varepsilon$ and, thus,

$$\begin{aligned} u - u^{\Delta} &= (u - u_{\varepsilon}) + (u_{\varepsilon} - u^{\Delta}) \\ &\leq C\varepsilon + \sup_{\bar{Q}_T \backslash \bar{Q}_{T-\Delta}} (u_{\varepsilon} - u^{\Delta})^{+} + C(T-t)\Delta\varepsilon^{-3} \\ &\leq \sup_{\bar{Q}_T \backslash \bar{Q}_{T-\Delta}} (u - u^{\Delta})^{+} + C(\varepsilon + \Delta\varepsilon^{-3}) \quad \text{in } \bar{Q}_T. \end{aligned}$$

By choosing $\varepsilon = \Delta^{\frac{1}{4}}$, we further obtain

$$u - u^{\Delta} \leq \sup_{\bar{Q}_T \backslash \bar{Q}_{T-\Delta}} (u - u^{\Delta})^{+} + C\Delta^{\frac{1}{4}} \quad \text{in } \bar{Q}_T.$$

We conclude using estimate (3.27) in Lemma 3.3.1. ■

3.3.2 Lower bound for the approximation error

To obtain a lower bound of $u - u^{\Delta}$, we cannot follow the above perturbation procedure to construct approximate smooth supersolutions to equation (3.1). This is because if we perturb its coefficients to obtain a viscosity supersolution, its convolution with the mollifier may no longer be a supersolution due to the convexity of equation (3.1) with respect to its terms, i.e. the opposite direction of Lemma 3.3.4 fails to hold (see Remark 2.4.14 also). Furthermore, interchanging the roles (as in [2] and [39]) of equation (3.1) and its approximation scheme (3.18) does not work either, because the solution u^{Δ} of the approximation scheme (and its perturbation solution) may in general lose the Hölder and Lipschitz continuity in (t,x). This is due to the lack of the continuous dependence result for the approximation scheme, compared with the continuous dependence result for (3.1) and its perturbation equation (3.28) (see Lemma A.0.1).

To overcome these difficulties, we follow the idea of Barles and Jakobsen [4] to build approximate supersolutions which are smooth at the "right points" by introducing an appropriate optimal switching stochastic control system. To apply this method to the problem herein, we first observe that, using the convex dual function L introduced in (3.7), we can write (3.1) as a HJB equation, namely,

$$\begin{cases} -\partial_t u + \sup_{q \in \mathbb{R}^d} \mathcal{L}^q \left(t, x, D_x u, D_x^2 u\right) = 0 & \text{in } Q_T; \\ u(T,x) = U(x) & \text{in } \mathbb{R}^d, \end{cases} \tag{3.33}$$

with

$$\mathcal{L}^q(t,x,p,X) := -\frac{1}{2}\mathrm{tr}\left(\sigma\sigma^T(t,x)X\right) - (b(t,x)-q)\cdot p - L(t,x,q).$$

It then follows from Proposition 3.2.3 (iv) that the supremum can be achieved at some point, say q^*, with $|q^*| \le \xi(|D_x u|)$. Furthermore, Proposition 3.2.2 implies that $|q^*| \le C$, for some constant C depending only on M and T. Thus, we rewrite the equation in (3.33) as

$$-\partial_t u + \sup_{q\in K} \mathcal{L}^q\left(t,x,D_x u, D_x^2 u\right) = 0, \tag{3.34}$$

where $K \subset \mathbb{R}^d$ is a compact set. Since K is separable, it has a countable dense subset, say $K_\infty = \{q_1, q_2, q_3, ...\}$ and, in turn, the continuity of $\mathcal{L}^q$ in q implies that

$$\sup_{q\in K} \mathcal{L}^q(t,x,p,X) = \sup_{q\in K_\infty} \mathcal{L}^q(t,x,p,X).$$

Therefore, the equation in (3.33) further reduces to

$$-\partial_t u + \sup_{q\in K_\infty} \mathcal{L}^q\left(t,x,D_x u, D_x^2 u\right) = 0.$$

For $m \ge 1$, we consider the approximations of (3.33),

$$\begin{cases} -\partial_t u^m + \sup\limits_{q\in K_m} \mathcal{L}^q\left(t,x,D_x u^m, D_x^2 u^m\right) = 0 & \text{in } Q_T; \\ u^m(T,x) = U(x) & \text{in } \mathbb{R}^d, \end{cases} \tag{3.35}$$

where $K_m := \{q_1, ..., q_m\} \subset K_\infty$, i.e. K_m consists of the first m points in K_∞ and satisfies $\cup_{m\ge 1} K_m = K_\infty$. It then follows from Proposition 2.1 of [4] that (3.35) admits a unique viscosity solution $u^m \in \mathcal{C}_b^1(\bar{Q}_T)$, with $|u^m|_1 \le C$. Furthermore, Arzela-Ascoli's theorem yields that there exists a subsequence of $\{u^m\}$, still denoted as $\{u^m\}$, such that, as $m \to \infty$,

$$u^m(t,x) \to u(t,x) \text{ locally uniformly in } (t,x) \in \bar{Q}_T. \tag{3.36}$$

Moreover, since $u(T,\cdot) = u^m(T,\cdot) = U$, it follows from their regularity result that, for $m \in \mathbb{Z}^+$

$$|u(t,\cdot) - u^m(t,\cdot)|_0 \le |u(t,\cdot) - u(T,\cdot)|_0 + |u^m(t,\cdot) - u^m(T,\cdot)|_0 \le C\sqrt{T-t}. \tag{3.37}$$

Next, we construct a sequence of (local) smooth supersolutions to approximate u^m. For this, we consider the optimal switching system

$$\begin{cases} \max\left\{-\partial_t v_i + \mathcal{L}^{q_i}(t,x,D_x v_i, D_x^2 v_i), v_i - \mathcal{M}_i^k v\right\} = 0 & \text{in } Q_T; \\ v_i(T,x) = U(x) & \text{in } \mathbb{R}^d, \end{cases} \tag{3.38}$$

where $i \in \mathcal{I} := \{1, ..., m\}$ and $\mathcal{M}_i^k v := \min_{j\ne i,\ j\in\mathcal{I}}\{v_j + k\}$, for some constant $k > 0$ representing the switching cost.

Proposition 3.3.6 *Suppose that Assumption 3.2.1 is satisfied. Then, there exists a unique viscosity solution $v = (v_1, \ldots, v_m)$ of the optimal switching* *system (3.38) such that $|v|_1 \leq C$. Moreover, for $i \in \mathcal{I}$,*

$$0 \leq v_i - u^m \leq C(k^{\frac{1}{3}} + k^{\frac{2}{3}}) \quad \text{in } \bar{Q}_T. \tag{3.39}$$

The proof essentially follows from Proposition 2.1 and Theorem 2.3 of [4] and it is thus omitted. We only remark that since we do not require the switching cost to satisfy $k \leq 1$, we keep the term $k^{\frac{2}{3}}$ in the above estimate. This will not affect the convergence rate of the approximation scheme.

Next, still following the approach of [4], we construct smooth approximations of v_i. Since in the continuation region of (3.38), the solution v_i satisfies the *linear* equation, namely,

$$-\partial_t v_i + \mathcal{L}^{q_i}(t, x, D_x v_i, D_x^2 v_i) = 0 \quad \text{in } \{(t,x) \in Q_T : v_i(t,x) < \mathcal{M}_i^k v(t,x)\},$$

we may perturb its coefficients to obtain a sequence of smooth supersolutions. This will in turn give a lower bound of the error $u^m - u^{\Delta}$. A subtle point herein is how to identify the continuation region by appropriately choosing the switching cost k. For this, we follow the idea used in Lemma 3.4 of [4].

Proposition 3.3.7 *Suppose that Assumption 3.2.1 holds. Let $u^{\Delta} \in \mathcal{C}_b(\bar{Q}_T)$ satisfy the approximation scheme (3.18) and $u^m \in \mathcal{C}_b^1(\bar{Q}_T)$ be the unique viscosity solution of the HJB equation (3.35). Then,*

$$u^{\Delta} - u^m \leq \sup_{\bar{Q}_T \setminus \bar{Q}_{T-\Delta}} (u^{\Delta} - u^m)^+ + C\Delta^{\frac{1}{10}} \quad \text{in } \bar{Q}_T.$$

Proof. Let $\varepsilon \in [0,1]$. In analogy to (3.28), we perturb the coefficients of the optimal switching system (3.38) and consider

$$\begin{cases} \max\left\{-\partial_t v_i^\varepsilon + \inf_{(\tau,e)\in\Theta^\varepsilon} \mathcal{L}^{q_i}(t+\tau, x+e, D_x v_i^\varepsilon, D_x^2 v_i^\varepsilon), v_i^\varepsilon - \mathcal{M}_i^k v^\varepsilon\right\} = 0 & \text{in } Q_{T+\varepsilon^2}; \\ v_i^\varepsilon(T+\varepsilon^2, x) = U(x) & \text{in } \mathbb{R}^d. \end{cases} \tag{3.40}$$

It then follows from Proposition 2.2 of [4] that (3.40) admits a unique viscosity solution, say $v^\varepsilon = (v_1^\varepsilon, \dots, v_m^\varepsilon)$, with $|v^\varepsilon|_1 \leq C$ and, moreover, for each $i \in \mathcal{I}$,

$$|v_i^\varepsilon - v_i| \leq C\varepsilon \quad \text{in } \bar{Q}_T, \tag{3.41}$$

where the constant C depends only on M and T. In turn, inequalities (3.39) and (3.41) imply that, for each $i \in \mathcal{I}$,

$$|v_i^\varepsilon - u^m| \leq |v_i^\varepsilon - v_i| + |v_i - u^m| \leq C(\varepsilon + k^{\frac{1}{3}} + k^{\frac{2}{3}}) \quad \text{in } \bar{Q}_T. \tag{3.42}$$

Next, we regularize v_i^ε by introducing $v_{i,\varepsilon}(t,x) := v_i^\varepsilon * \rho_\varepsilon(t,x)$, for $(t,x) \in \bar{Q}_T$, where ρ_ε is the mollifier defined in (3.30). Then, $v_{i,\varepsilon} \in \mathcal{C}_b^\infty(\bar{Q}_T)$,

$$|v_{i,\varepsilon} - v_i^\varepsilon|_0 \leq C\varepsilon, \tag{3.43}$$

and, moreover, for positive integers m and multi-index n,

$$|\partial_t^m D_x^n v_{i,\varepsilon}|_0 \leq C\varepsilon^{1-2m-|n|}. \tag{3.44}$$

We introduce the function $w_\varepsilon := \min_{i\in\mathcal{I}} v_{i,\varepsilon}$, which is smooth in $\bar{Q}_T$ except for finitely many points. Then, (3.42) and (3.43) yield

$$|u^m - w_\varepsilon| \le C(\varepsilon + k^{\frac{1}{3}} + k^{\frac{2}{3}}) \ \text{ in } \bar{Q}_T. \tag{3.45}$$

For each $(t,x)\in\bar{Q}_T$, let $j := \arg\min_{i\in\mathcal{I}} v_{i,\varepsilon}(t,x)$. Then, $w_\varepsilon(t,x) = v_{j,\varepsilon}(t,x)$ and, for such j, we obtain that

$$v_{j,\varepsilon}(t,x) - \mathcal{M}_j^k v_\varepsilon(t,x) = \max_{i\neq j, i\in\mathcal{I}} \{v_{j,\varepsilon}(t,x) - v_{i,\varepsilon}(t,x) - k\} \le -k.$$

In turn, inequality (3.43) implies that

$$v_j^\varepsilon(t,x) - \mathcal{M}_j^k v^\varepsilon(t,x) \le v_{j,\varepsilon}(t,x) - \mathcal{M}_j^k v_\varepsilon(t,x) + C\varepsilon \le -k + C\varepsilon.$$

Furthermore, since $|v^\varepsilon|_1 \le C$ for $v^\varepsilon = (v_1^\varepsilon, \dots, v_m^\varepsilon)$, we also have

$$\begin{aligned} v_j^\varepsilon(t-\tau,x-e) - \mathcal{M}_j^k v^\varepsilon(t-\tau,x-e) &\le v_j^\varepsilon(t,x) - \mathcal{M}_j^k v^\varepsilon(t,x) + C(|\tau|^{\frac{1}{2}} + |e|) \\ &\le -k + C\varepsilon + 2C\varepsilon, \end{aligned}$$

for any $(\tau,e)\in\Theta^\varepsilon$. If we then choose $k = 4C\varepsilon$, we obtain that, for any $(\tau,e)\in\Theta^\varepsilon$,

$$v_j^\varepsilon(t-\tau,x-e) - \mathcal{M}_j^k v^\varepsilon(t-\tau,x-e) < 0.$$

Therefore, the point $(t-\tau,x-e)$, for $(\tau,e)\in\Theta^\varepsilon$, is in the continuation region of (3.40). Thus,

$$-\partial_t v_j^\varepsilon(t-\tau,x-e) + \mathcal{L}^{q_j}\left(t,x,D_x v_j^\varepsilon(t-\tau,x-e), D_x^2 v_j^\varepsilon(t-\tau,x-e)\right) \ge 0.$$

Using the definition of $v_{j,\varepsilon}$ and that $\mathcal{L}^{q_j}$ is linear in $D_x v_j^\varepsilon$ and $D_x^2 v_j^\varepsilon$, we further have

$$\begin{aligned} &- \partial_t v_{j,\varepsilon}(t,x) + \mathcal{L}^{q_j}\left(t,x,D_x v_{j,\varepsilon}(t,x), D_x^2 v_{j,\varepsilon}(t,x)\right) \\ &= \int_{-\varepsilon^2<\tau<0}\int_{|e|<\varepsilon} \left(-\partial_t v_j^\varepsilon(t-\tau,x-e) + \mathcal{L}^{q_j}\left(t,x,D_x v_j^\varepsilon(t-\tau,x-e), D_x^2 v_j^\varepsilon(t-\tau,x-e)\right)\right) \\ &\quad \times \rho_\varepsilon(\tau,e) de d\tau \ge 0. \end{aligned} \tag{3.46}$$

Next, we observe that, for $(t,x)\in\bar{Q}_{T-\Delta}$, the definition of j implies that $w_\varepsilon(t,x) = v_{j,\varepsilon}(t,x)$ and $w_\varepsilon(t+\Delta,\cdot) \le v_{j,\varepsilon}(t+\Delta,\cdot)$. Then, applying Proposition 3.2.7 (i) (iii) and estimate (3.44), we obtain, for any $(t,x)\in\bar{Q}_{T-\Delta}$,

$$\begin{aligned} &S(\Delta,t,x,w_\varepsilon(t,x),w_\varepsilon(t+\Delta,\cdot)) \\ &\ge S(\Delta,t,x,v_{j,\varepsilon}(t,x),v_{j,\varepsilon}(t+\Delta,\cdot)) \\ &\ge \ - \partial_t v_{j,\varepsilon}(t,x) + \sup_{q\in\mathbb{R}^d} \mathcal{L}^q(t,x,D_x v_{j,\varepsilon}(t,x), D_x^2 v_{j,\varepsilon}(t,x)) - C\Delta\varepsilon^{-3} \\ &\ge \ - \partial_t v_{j,\varepsilon}(t,x) + \mathcal{L}^{q_j}(t,x,D_x v_{j,\varepsilon}(t,x), D_x^2 v_{j,\varepsilon}(t,x)) - C\Delta\varepsilon^{-3} \ge -C\Delta\varepsilon^{-3}, \end{aligned}$$

for some constant C depending only on M and T, where we used (3.46) in the last inequality. In turn, the comparison result in Proposition 3.2.8 implies that

$$u^\Delta - w_\varepsilon \le \sup_{\bar{Q}_T\setminus\bar{Q}_{T-\Delta}} (u^\Delta - w_\varepsilon)^+ + C(T-t)\Delta\varepsilon^{-3} \ \text{ in } \bar{Q}_T.$$

Combining the above inequality with (3.45), we further get

$$
\begin{aligned}
u^\Delta - u^m &= (u^\Delta - w_\varepsilon) + (w_\varepsilon - u^m) \\
&\leq \sup_{\bar{Q}_T \setminus \bar{Q}_{T-\Delta}} (u^\Delta - w_\varepsilon)^+ + C(T-t)\Delta\varepsilon^{-3} + C(\varepsilon + k^{\frac{1}{3}} + k^{\frac{2}{3}}) \\
&\leq \sup_{\bar{Q}_T \setminus \bar{Q}_{T-\Delta}} (u^\Delta - u^m)^+ + C(\varepsilon + \varepsilon^{\frac{1}{3}} + \varepsilon^{\frac{2}{3}} + \Delta\varepsilon^{-3}) \\
&\leq \sup_{\bar{Q}_T \setminus \bar{Q}_{T-\Delta}} (u^\Delta - u^m)^+ + C\Delta^{\frac{1}{10}} \quad \text{in } \bar{Q}_T,
\end{aligned}
$$

where we used $k = 4C\varepsilon$ in the second to last inequality, and chose $\varepsilon = \Delta^{\frac{3}{10}}$ in the last inequality. ■

We are now ready to establish a lower bound for the approximation error.

Theorem 3.3.8 *Suppose that Assumption 3.2.1 holds. Let $u^\Delta \in \mathcal{C}_b(\bar{Q}_T)$ satisfy the approximation scheme (3.18) and $u \in \mathcal{C}_b^1(\bar{Q}_T)$ be the unique viscosity solution of equation (3.1). Then,*

$$
u - u^\Delta \geq -C\Delta^{\frac{1}{10}} \quad \textit{in } \bar{Q}_T.
$$

Proof. Applying Proposition 3.3.7 to the sequence $\{u^m\}$, we get for $m \in \mathbb{Z}^+$,

$$
\begin{aligned}
u^\Delta - u &= (u^\Delta - u^m) + (u^m - u) \\
&\leq \sup_{\bar{Q}_T \setminus \bar{Q}_{T-\Delta}} (u^\Delta - u^m)^+ + C\Delta^{\frac{1}{10}} + (u^m - u) \\
&\leq \sup_{\bar{Q}_T \setminus \bar{Q}_{T-\Delta}} (u^\Delta - u)^+ + \sup_{\bar{Q}_T \setminus \bar{Q}_{T-\Delta}} (u - u^m)^+ + C\Delta^{\frac{1}{10}} + u^m - u \\
&\leq C\Delta^{\frac{1}{10}} + u^m - u,
\end{aligned}
$$

where we used estimate (3.27) and (3.37) in the last inequality. Sending $m \to \infty$ and using (3.36), we conclude. ■

3.4 A numerical example

In this section we present a numerical result, applying the approximation scheme (3.18) for the case

$$
\sigma(t,x) = 1,\ b(t,x) = 0,\ H(t,x,p) = p^2/2,\ T = 1,
$$

and $U(x) = 0 \vee x \wedge K$ in the semilinear PDE (3.1). Then the equation in (3.1) becomes the Cole-Hopf equation (see [26]):

$$
-\partial_t u(t,x) - \frac{1}{2}\partial_{xx} u(t,x) + \frac{1}{2}(\partial_x u(t,x))^2 = 0. \tag{3.47}
$$

It is well known that, by the Cole-Hopf transformation (see [26]), the function $v(t,x) := e^{-u(t,x)}$ satisfies the standard heat equation

$$\partial_t v(t,x) + \frac{1}{2}\partial_{xx} v(t,x) = 0,$$

with $v(T,x) = e^{-U(x)} = e^{-0\vee x\wedge K}$. In turn,

$$\begin{aligned} v(t,x) &= \Phi(-\frac{x}{\sqrt{T-t}}) + e^{-x+(T-t)/2}\left(\Phi(\frac{K-x+T-t}{\sqrt{T-t}}) - \Phi(\frac{-x+T-t}{\sqrt{T-t}})\right) \\ &\quad + e^{-K}\Phi(-\frac{K-x}{\sqrt{T-t}}), \end{aligned}$$

where Φ is the standard normal cumulative distribution function and, thus, we obtain the explicit solution

$$u(t,x) = -\log v(t,x).$$

We use this exact solution as a benchmark and compare it with the approximate solution obtained by the monotone scheme (3.18). Moreover, we also compare our results with the ones obtained via the standard Howard's finite difference (FD) algorithm (see, for example, [8] for a detailed discussion of Howard's FD scheme).

To make the monotone scheme (3.18) implementable, we further discretize the space variable and since Howard's scheme is based on finite difference method, for the comparison purpose we also numerically compute the conditional expectation appearing in the backward operator $\mathbf{S}_t(\Delta)$ (cf. (3.8)) using the finite difference method. However, we emphasize that, different from Howard's scheme, the splitting approximation itself does not depend on finite difference method, as long as one can find an efficient way to compute conditional expectations. For example, one may use Gauss-Hermite quadrature approximation as in [56, 57], although this may be a source of exponential computational complexity in the dimension d.

To numerically compute the finite dimension minimization problem in the backward operator $\mathbf{S}_t(\Delta)$ (cf. (3.8)), since the finite difference method already provides us with all the partition points in x, we further discretise the partition with a multiplier of 10 and use the simple brute force method to find the minimizers within those points and the corresponding minimal values. We note that the simple brute force method makes the algorithm have a quadratic computational complexity in the number of mesh points of space discretization, which is a main disadvantage of the monotone scheme compared to finite difference schemes who typically has a linear complexity. Moreover, in the above example (3.47), the Legendre transform has a simple closed-form $L(t,x,q) = q^2/2$ whereas in the general cases where the closed-form formula is not available, numerical optimization methods such as gradient descent and Nelder-Mean simplex algorithm need to be applied to compute the Legendre transform appearing in the backward operator $\mathbf{S}_t(\Delta)$.

Figures 3.1 and 3.2 demonstrate the performance of the monotone scheme (3.18) with the parameter $K = 5$. They illustrate how the approximation solutions converge as we increase the number of time steps T/Δ. For our parameter values, $\Delta = 0.1$ (so $T/\Delta = 10$) is sufficient for the values to converge, as the relative error is already negligible (0.056%).

Figure 3.3 compares the values numerically computed by the monotone scheme and the Howard's FD scheme with different time steps. It shows that the monotone scheme gives a much better approximation than the Howard's FD scheme does throughout all the time steps. In particular, we observe that when the time step $\Delta = 0.1$ (so $T/\Delta = 10$), the numerical solution computed by our approximation scheme is far more accurate than the one computed by the FD scheme. The relative error is 0.056% for the former and 0.142% for the latter. It also shows the monotone scheme converges linearly with time step Δ, and this is consistent with our theoretical result in Theorem 3.3.2. Table 3.1 further compares the computation errors and costs between the approximation scheme (3.18) and the Howard's FD scheme. Since there involves an additional minimization step in the approximation scheme (3.18), its computation costs are higher than the FD scheme. However, we observe that when the time step is small (e.g. $\Delta = 0.1$), the computation times for both schemes are extremely fast (less than 0.05 second).

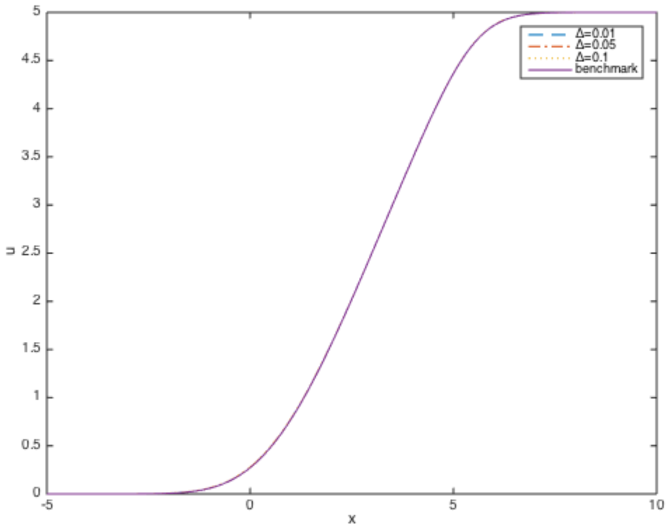

Figure 3.1: Approximation for $u(0, x)$ with various time steps $\Delta = 0.01/0.05/0.1$.

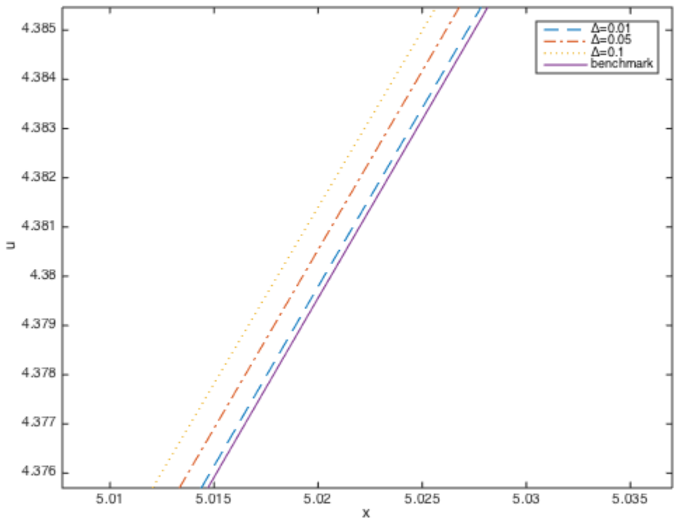

Figure 3.2: Approximation for $u(0, x)$ with various time steps $\Delta = 0.01/0.05/0.1$. This figure zooms in on Figure 3.1 at $x \approx 5$.

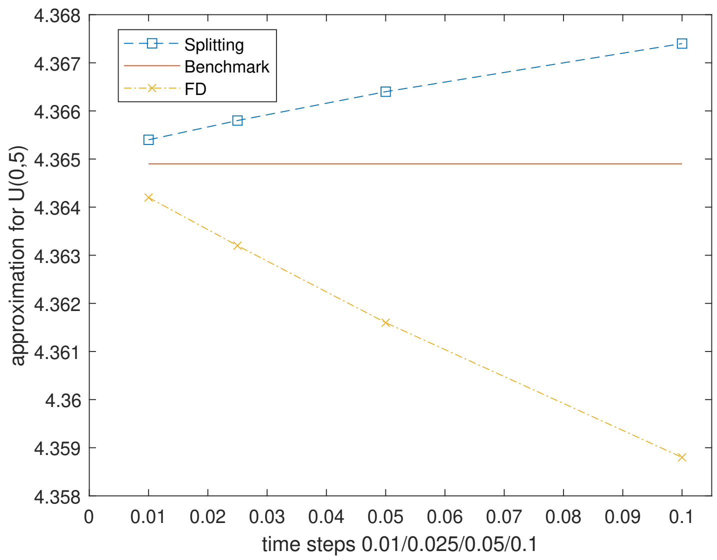

Figure 3.3: Comparison of exact value and approximate values for $u(0,5)$ via the monotone scheme (3.18) and the Howard's FD scheme with various time steps $\Delta = 0.01/0.025/0.05/0.1$.

Time steps / Numerical schemes	0.01	0.025	0.05	0.1
splitting approx. value	4.3655	4.3658	4.3664	4.3674
approx. error	0.012%	0.02%	0.032%	0.056%
running time (in seconds)	18.78	1.07	0.16	0.04
FD approx. value	4.3642	4.3632	4.3616	4.3588
approx. error	0.016%	0.039%	0.076%	0.142%
running time (in seconds)	7.01	0.43	0.03	0.01

Table 3.1: Comparison of running errors and costs for approximating $u(0,5)$ via the approximation scheme (3.18) and the Howard's FD scheme with various time steps $\Delta = 0.01/0.025/0.05/0.1$.

Chapter 4

A monotone approximation scheme for variational inequalities with convex and coercive Hamiltonians

In this chapter, we extend the work in Chapter 3 and consider variational versions of problem (3.1). These are naturally related to optimal stopping and to singular stochastic optimization problems, both directly related to various applications with early-exercise, fixed, and/or proportional transaction costs, irreversible investment decisions, etc. Specifically, we consider semilinear parabolic variational inequalities of the form

$$\begin{cases} \max\{-\partial_t u + g(t,x,D_x u, D_x^2 u), u - f(t,x)\} = 0 & \text{in } Q_T; \\ u(T,x) = U(x) & \text{in } \mathbb{R}^d, \end{cases} \tag{4.1}$$

where

$$g(t,x,p,X) := -\frac{1}{2}\mathrm{tr}\left(\sigma\sigma^T(t,x)X\right) - b(t,x)\cdot p + H(t,x,p).$$

In comparison with the semilinear parabolic PDE (3.1) considered in Chapter 3, an obstacle terms f is added in equation (4.1). Note that in the continuation region where $u < f$, equation (4.1) reduces exactly to (3.1).

We propose an adapted approximation scheme for variational inequality (4.1) by using a natural extension of scheme (3.18) proposed for (3.1) in Chapter 3:

$$\max\left\{S(\Delta,t,x,u^\Delta(t,x),u^\Delta(t+\Delta,\cdot)), u^\Delta(t,x) - f(t,x)\right\} = 0,$$

where u^Δ is the solution of the scheme used to approximate the (viscosity) solution of (4.1). The formal introduction of the approximation scheme is given in (4.3) in the following section.

Similar to Chapter 3, we next establish the convergence rate of the scheme solution u^Δ to the unique (viscosity) solution u of (4.1) by obtaining the error estimates $u - u^\Delta$. We apply again the *shaking coefficients technique* (for upper bound) and the *optimal switching approximation* (for lower bound) with the newly added obstacle term $f(t,x)$.

The lower bound of $u - u^\Delta$ is the main difficulty of this chapter. This is because the variational inequality (4.1), which incorporates the obstacle $f(t,x)$, can not be approximated by a standard optimal switching system of the form (3.38). Instead, we modify the standard optimal switching system and introduce a obstacle type of switching system, which can be proved to approximate HJB variational inequalities. We give the well-posedness, regularity as well as continuous dependence result for this *obstacle switching system* in Section 4.3. A similar type of obstacle switching system is used in Dumitrescu, Reisinger and Zhang [23], where they designed a numerical scheme directly based on such type of switching systems to approximate a class of HJB variational inequalities with jumps, but proved convergence only (without any convergence rate). Furthermore, they proved the existence of viscosity solutions of the obstacle switching system by constructing the solution directly from the numerical solutions. In contrast, we prove the existence by showing that viscosity solutions of the obstacle switching system can be represented as the value functions of mixed optimal stopping and switching problems in Proposition 4.3.2.

A highly related work to this chapter is Jakobsen [39], where he obtained error bounds for general monotone approximation schemes for Bellman equations arising in a stochastic optimal stopping and control problem. Due to the convex and coercive property of the Hamiltonian $H(t,x,p)$, (4.1) can be written as the same type of Bellman equation, but with control set and coefficients unbounded. This is not the case in [39]. Furthermore, he used the same *shaking coefficients technique* to derive an error bound for one side, but for the other side he interchanged the roles of the approximation scheme and the original equation, based on an additional assumption (Assumption 2.5 in [39]) that the scheme solution has enough regularity. Unfortunately, our proposed scheme does not satisfy this assumption and thus, we need to apply the optimal switching approximation technique which gives a slower convergence rate for lower bound. Finally, a common point between his work and ours is that the scheme solution u^{Δ} is defined at every point in $\bar{Q}_T$ rather than some certain time (and space) grids.

This chapter is organized as follows. In the next section we introduce the approximation scheme for (4.1) with some of its properties. In Section 4.2, we prove its convergence and establish the convergence rate by obtaining the upper and lower bounds for the error estimate. Section 4.3 is then devoted to introducing the obstacle switching systems. Some technical proofs are provided in the appendix.

4.1 The adapted approximation scheme for variational inequality (4.1)

In this section, we adapt the approximation scheme (3.18) to obtain a scheme for solving variational inequality (4.1). We suppose the Assumption 3.2.1 still holds in this chapter. In addition, we assume the following condition for the obstacle term f.

Assumption 4.1.1 *The real-valued obstacle f has finite norm $|f|_1 \leq M$ for the same $M > 0$ in Assumption 3.2.1(i). Moreover, $f(T,\cdot) \geq U$ in $\mathbb{R}^d$.*

Unless state otherwise, in this chapter we still denote by $C := C(T, M)$ some constant that depends only on T and M. Under the Assumption 3.2.1 and 4.1.1, we have the same existence, uniqueness and regularity results for the solution of variational inequality (4.1). The proof requires some modification of the proofs of Proposition 3.2.2 and 3.3.3 stated in Appendix A to incorporate the obstacle term f, and is provided in Appendix B.

Proposition 4.1.2 *Suppose that Assumption 3.2.1 and 4.1.1 are satisfied. Then, there exists a unique viscosity solution $u \in \mathcal{C}_b^1(\bar{Q}_T)$ of (4.1), with $|u|_1 \leq C$.*

Based on the backward operator $\mathbf{S}_t(\Delta)$ and the scheme (3.18) constructed in Chapter 3, we now introduce the adapted approximation scheme for solving the solution u of variational inequality (4.1): For $\Delta \in (0, T)$ and $(t, x) \in \bar{Q}_{T-\Delta}$, we apply the adapted iterative algorithm

$$u^{\Delta}(t, x) = \min\left\{\mathbf{S}_t(\Delta)u^{\Delta}(t + \Delta, \cdot)(x), f(t, x)\right\} \tag{4.2}$$

with $u^{\Delta}(T, \cdot) = U(\cdot)$ and $\mathbf{S}_t(\Delta)$ defined in (3.8). The values of $u^{\Delta}(t, x)$ when $T - \Delta < t < T$ are again obtained by a standard linear interpolation.

Specifically, the adapted approximation scheme is given by

$$\begin{cases} \bar{S}(\Delta, t, x, u^{\Delta}(t, x), u^{\Delta}(t + \Delta, \cdot)) = 0 & \text{in } \bar{Q}_{T-\Delta}, \\ u^{\Delta}(t, x) = g^{\Delta}(t, x) & \text{in } \bar{Q}_T \backslash \bar{Q}_{T-\Delta}, \end{cases} \tag{4.3}$$

where $\bar{S} : (0, T) \times \bar{Q}_{T-\Delta} \times \mathbb{R} \times \mathcal{C}_b(\mathbb{R}^d) \to \mathbb{R}$, and $g^{\Delta} : \bar{Q}_T \backslash \bar{Q}_{T-\Delta} \to \mathbb{R}$ are defined respectively by

$$\bar{S}(\Delta, t, x, p, v) = \max\{S(\Delta, t, x, p, v), p - f(t, x)\}, \tag{4.4}$$

and

$$g^{\Delta}(t, x) = \omega_1(t)U(x) + \omega_2(t)\min\{\mathbf{S}_{T-\Delta}(\Delta)U(x), f(T - \Delta, x)\}, \tag{4.5}$$

with $\omega_1(t) = (t + \Delta - T)/\Delta$ and $\omega_2(t) = (T - t)/\Delta$ being the linear interpolation weights, and S defined in (3.19).

The stability, monotonicity and consistency property of the adapted approximation scheme (4.3) still hold.

Proposition 4.1.3 *Suppose that Assumption 3.2.1 and 4.1.1 are satisfied. Then,*

(i) (Stability) The approximation scheme (4.3) admits a unique solution $u^{\Delta} \in \mathcal{C}_b(\bar{Q}_T)$ with $|u^{\Delta}|_0 \leq C$.

(ii) (Monotonicity) For any $p, c_1, c_2 \in \mathbb{R}$, and any function $u, v \in \mathcal{C}_b(\mathbb{R}^d)$ with $u \leq v$,

$$\bar{S}(\Delta, t, x, p + c_1, u + c_2) \geq \bar{S}(\Delta, t, x, p, v) + \min\{\frac{c_1 - c_2}{\Delta}, c_1\} \quad \text{in } \bar{Q}_{T-\Delta}.$$

(iii) (Consistency) For any $\phi \in \mathcal{C}_b^{\infty}(\bar{Q}_T)$,

$$\begin{aligned} &|\max\{-\partial_t\phi + g(t, x, D_x\phi, D_x^2\phi), \phi - f(t, x)\} - \bar{S}(\Delta, t, x, \phi, \phi(t + \Delta, \cdot))| \\ &\leq C\Delta\left(|\partial_t^2\phi|_0 + |D_x^4\phi|_0 + |\partial_t D_x^2\phi|_0 + \mathcal{R}(\phi)\right) \quad \text{in } \bar{Q}_{T-\Delta}, \end{aligned} \tag{4.6}$$

where the constant C depends only on $[\phi]_{2,1}$, M and T, and $\mathcal{R}(\phi)$ represents the "insignificant" terms containing the lower order derivatives of ϕ.

Proof. (i) By the stability property (iv) of Proposition 3.2.5, we know that

$$|\mathbf{S}_t(\Delta)u^\Delta(t+\Delta,\cdot)|_0 \le C\Delta + |u^\Delta(t+\Delta,\cdot)|_0.$$

Since $|f|_0 \le M$, we have from (4.2) that for $0 \le t \le T-\Delta$,

$$|u^\Delta(t,\cdot)|_0 \le \max\{C\Delta + |u^\Delta(t+\Delta,\cdot)|_0, M\} \le C\Delta + \max\{|u^\Delta(t+\Delta,\cdot)|_0, M\},$$

and thus

$$\max\{|u^\Delta(t,\cdot)|_0, M\} \le C\Delta + \max\{|u^\Delta(t+\Delta,\cdot)|_0, M\},$$

where $C = \max\{|L^*(0)|, |H^*(0)|\}$. By backward induction and the definition of g^Δ in (4.3), we conclude that

$$|u^\Delta|_0 \le CT + \max\{\sup_{t\in(T-\Delta,T]} |g^\Delta(t,\cdot)|_0, M\} \le C.$$

(ii) It follows from the monotonicity property of S, Proposition 3.2.7(i), that

$$\begin{aligned}
&\bar{S}(\Delta,t,x,p+c_1,u+c_2)\\
=\ &\max\{S(\Delta,t,x,p+c_1,u+c_2), p+c_1-f(t,x)\}\\
\ge\ &\max\{S(\Delta,t,x,p,v)+\frac{c_1-c_2}{\Delta}, p+c_1-f(t,x)\}\\
\ge\ &\bar{S}(\Delta,t,x,p,v)+\min\{\frac{c_1-c_2}{\Delta}, c_1\}.
\end{aligned}$$

(iii) follows directly from Proposition 3.2.7(iii) and the fact that $h(x) := \max\{x,c\}$ is Lipschitz in x for any constant c. ■

The monotonicity property (i) in Proposition 4.1.3 then implies the following comparison result for the approximation scheme (4.3), which will be used throughout this paper. The proof is analogous to Proposition 3.2.8, with a slight difference to accommodate the extension to the variational inequality case.

Proposition 4.1.4 *Suppose that Assumption 3.2.1 and 4.1.1 are satisfied, and that u, $v \in \mathcal{C}_b(\bar{Q}_T)$ are such that*

$$\bar{S}(\Delta,t,x,u,u(t+\Delta,\cdot)) \le h_1 \text{ in } \bar{Q}_{T-\Delta},$$

$$\bar{S}(\Delta,t,x,v,v(t+\Delta,\cdot)) \ge h_2 \text{ in } \bar{Q}_{T-\Delta},$$

for some h_1, $h_2 \in \mathcal{C}_b(\bar{Q}_{T-\Delta})$. Then,

$$u - v \le \sup_{\bar{Q}_T\setminus\bar{Q}_{T-\Delta}} (u-v)^+ + (T-t+1)\sup_{\bar{Q}_{T-\Delta}} (h_1-h_2)^+ \text{ in } \bar{Q}_T. \tag{4.7}$$

Proof. The inequality (4.7) holds obviously in $\bar{Q}_T\setminus\bar{Q}_{T-\Delta}$. We then prove it holds in $\bar{Q}_{T-\Delta}$ by induction. To this end, suppose (4.7) holds in $\{t+\Delta\}\times\mathbb{R}^d$

for some $t+\Delta$ such that $t \le T-\Delta$, then since $\bar{S}(\Delta, t, x, u, u(t+\Delta,\cdot)) \le h_1$, we have

a) $u(t,x) \le \mathbf{S}_t(\Delta)u(t+\Delta,\cdot)(x) + \Delta h_1(t,x)$ and *b)* $u(t,x) - f(t,x) \le h_1(t,x)$,

and similarly,

c) $v(t,x) \ge \mathbf{S}_t(\Delta)v(t+\Delta,\cdot)(x) + \Delta h_2(t,x)$ or *d)* $v(t,x) - f(t,x) \ge h_2(t,x)$.

Suppose c) holds, then using a) and Proposition 3.2.5(ii), we have

$$\begin{aligned}
&(u-v)(t,x) \\
\le\ & \sup_{x\in\mathbb{R}^d}(u-v)(t+\Delta,x) + \Delta(h_1-h_2)(t,x) \\
\le\ & \sup_{x\in\mathbb{R}^d}(u-v)(t+\Delta,x) + \Delta \sup_{\bar{Q}_{T-\Delta}}(h_1-h_2)^+ \\
\le\ & \sup_{\bar{Q}_T\backslash\bar{Q}_{T-\Delta}}(u-v)^+ + (T-t-\Delta+1)\sup_{\bar{Q}_{T-\Delta}}(h_1-h_2)^+ + \Delta\sup_{\bar{Q}_{T-\Delta}}(h_1-h_2)^+ \\
=\ & \sup_{\bar{Q}_T\backslash\bar{Q}_{T-\Delta}}(u-v)^+ + (T-t+1)\sup_{\bar{Q}_{T-\Delta}}(h_1-h_2)^+.
\end{aligned}$$

Suppose d) holds, then together with b), we have

$$(u-v)(t,x) \le (h_1-h_2)(t,x) \le \sup_{\bar{Q}_T\backslash\bar{Q}_{T-\Delta}}(u-v)^+ + (T-t+1)\sup_{\bar{Q}_{T-\Delta}}(h_1-h_2)^+.$$

Thus, in either case, (4.7) holds in $\{t\}\times\mathbb{R}^d$, and this completes the proof. ∎

4.2 Convergence rate of the approximation scheme

In this section, we establish the convergence rate of the approximate solution u^Δ to the viscosity solution u of the variational inequality (4.1). Similar to Section 3.3, we first derive the convergence rate in the final time interval $\bar{Q}_T\backslash\bar{Q}_{T-\Delta}$, and then establish the upper and lower bound of approximation error $u-u^\Delta$ within the whole domain $\bar{Q}_T$ separately.

Lemma 4.2.1 *Suppose that Assumption 3.2.1 and 4.1.1 are satisfied. Let $u^\Delta \in \mathcal{C}_b(\bar{Q}_T)$ be the unique solution of the approximation scheme (4.3) and $u \in \mathcal{C}_b^1(\bar{Q}_T)$ be the unique viscosity solution of equation (4.1). Then,*

$$\sup_{\bar{Q}_T\backslash\bar{Q}_{T-\Delta}}|u-u^\Delta| \le C\sqrt{\Delta}. \tag{4.8}$$

Proof. From Proposition 3.2.5(v), we have $|U-\mathbf{S}_{T-\Delta}(\Delta)U|_0 \le C\sqrt{\Delta}$. On the other hand, by Assumption 4.1.1, for any $x\in\mathbb{R}^d$, $f(T-\Delta,x) \ge f(T,x) - [f]_{\mathcal{C}^{1/2,1}}\sqrt{\Delta} \ge U(x) - M\sqrt{\Delta}$. The above two inequalities then imply that for any $x\in\mathbb{R}^d$,

$$|U(x) - \min\{\mathbf{S}_{T-\Delta}(\Delta)U(x), f(T-\Delta,x)\}| \le C\sqrt{\Delta}. \tag{4.9}$$

Then from (4.3), we have, for $(t,x) \in \bar{Q}_T \backslash \bar{Q}_{T-\Delta}$,

$$
\begin{aligned}
&|u(t,x) - u^{\Delta}(t,x)| = |u(t,x) - g^{\Delta}(t,x)| \\
&= |u(t,x) - u(T,x) + \omega_2(t)(U(x) - \min\{\mathbf{S}_{T-\Delta}(\Delta)U(x), f(T-\Delta,x)\})| \\
&\le |u(t,x) - u(T,x)| + |U(x) - \min\{\mathbf{S}_{T-\Delta}(\Delta)U(x), f(T-\Delta,x)\}| \\
&\le C(\sqrt{|T-t|} + \sqrt{\Delta}) \le C\sqrt{\Delta},
\end{aligned}
$$

where the second to last inequality follows from the regularity property of the solution u (cf. Proposition 4.1.2) and (4.9). ■

The following convergence rate result is straightforward. Its proof is essentially the same as that of Theorem 3.3.2 and thus we omit it.

Theorem 4.2.2 *Suppose that Assumption 3.2.1 and 4.1.1 are satisfied, and that equation (4.1) admits a unique smooth solution $u \in \mathcal{C}_b^{\infty}(\bar{Q}_T)$. Then,*

$$
|u - u^{\Delta}| \le C\Delta \quad \text{in } \bar{Q}_T.
$$

4.2.1 Upper bound for the approximation error

We now apply the *shaking coefficients technique* again to derive an upper bound for the approximation error within the whole domain $\bar{Q}_T$ for the general $u \in \mathcal{C}_b^1(\bar{Q}_T)$ case.

To this end, for small enough $\varepsilon \ge 0$, we extend the functions f and $\eta := \sigma, b$ to $Q_{T+\varepsilon^2}^{-\varepsilon^2} := [-\varepsilon^2, T+\varepsilon^2) \times \mathbb{R}^d$ and H to $Q_{T+\varepsilon^2}^{-\varepsilon^2} \times \mathbb{R}^d$, so that Assumption 3.2.1 and 4.1.1 still hold. We then define $\eta^{\theta}(t,x) := \eta(t+\tau, x+e)$ and $H^{\theta}(t,x,p) := H(t+\tau, x+e, p)$, where $\theta = (\tau, e)$ with $\theta \in \Theta^{\varepsilon} := [-\varepsilon^2, 0] \times \varepsilon B(0,1)$. We then consider the perturbed version of (4.1), namely,

$$
\begin{cases}
\max\{-\partial_t u^{\varepsilon} + \sup\limits_{\theta \in \Theta^{\varepsilon}} g^{\theta}(t,x,D_x u^{\varepsilon}, D_x^2 u^{\varepsilon}), u^{\varepsilon} - f(t,x)\} = 0 & \text{in } Q_{T+\varepsilon^2}; \\
u^{\varepsilon}(T+\varepsilon^2, x) = U(x) & \text{in } \mathbb{R}^d.
\end{cases}
\tag{4.10}
$$

where

$$
g^{\theta}(t,x,p,X) = -\frac{1}{2}\mathrm{Trace}\left(\sigma^{\theta}{\sigma^{\theta}}^T(t,x)X\right) - b^{\theta}(t,x)\cdot p + H^{\theta}(t,x,p).
$$

Note that when the perturbation parameter $\varepsilon = 0$, equations (4.10) and (4.1) coincide.

We establish existence, uniqueness and regularity results for the perturbed equation (4.10), and a comparison between u and u^{ε}. Their proofs are provided in Appendix B.

Proposition 4.2.3 *Suppose that Assumption 3.2.1 and 4.1.1 are satisfied. Then, there exists a unique viscosity solution $u^{\varepsilon} \in \mathcal{C}_b^1(\bar{Q}_{T+\varepsilon^2})$ of (4.10), with $|u^{\varepsilon}|_1 \le C$. Moreover,*

$$
|u - u^{\varepsilon}| \le C\varepsilon \quad \text{in } \bar{Q}_T. \tag{4.11}
$$

Next, we regularize u^{ε} and define

$$
u_{\varepsilon}(t,x) = u^{\varepsilon} * \rho_{\varepsilon}(t,x) = \int_{-\varepsilon^2<\tau<0}\int_{|e|<\varepsilon} u^{\varepsilon}(t-\tau, x-e)\rho_{\varepsilon}(\tau,e)\,de\,d\tau,
$$

where ρ_ε is defined in (3.30), and obtain again that $u_\varepsilon \in \mathcal{C}_b^\infty(\bar{Q}_T)$,

$$|u^\varepsilon - u_\varepsilon|_0 \le C\varepsilon, \tag{4.12}$$

and

$$|\partial_t^i D_x^j u_\varepsilon|_0 \le C\varepsilon^{1-2i-|j|}. \tag{4.13}$$

We observe from (4.10) that the function $u^\varepsilon(t-\tau, x-e)$ is a viscosity subsolution of equation (3.1) in Q_T for any $\theta \in \Theta^\varepsilon$. Following the same arguments in the proof of Lemma 3.3.4, we deduce that $u_\varepsilon(t,x)$ is still a subsolution of (3.1) in Q_T, namely,

$$-\partial_t u_\varepsilon + g(t, x, D_x u_\varepsilon, D_x^2 u_\varepsilon) \le 0. \tag{4.14}$$

Theorem 4.2.4 *Suppose that Assumption 3.2.1 and 4.1.1 are satisfied. Let $u^\Delta \in \mathcal{C}_b(\bar{Q}_T)$ be the unique solution of the approximation scheme (4.3) and $u \in \mathcal{C}_b^1(\bar{Q}_T)$ be the unique viscosity solution of equation (4.1). Then,*

$$u - u^\Delta \le C\Delta^{\frac{1}{4}} \text{ in } \bar{Q}_T.$$

Proof. From (4.10), $u^\varepsilon - f \le 0$ in $Q_{T+\varepsilon^2}$. This yields that for $(t,x) \in \bar{Q}_{T-\Delta}$,

$$\begin{aligned} u_\varepsilon(t,x) - f(t,x) &\le \int_{-\varepsilon^2<\tau<0} \int_{|e|<\varepsilon} (f(t-\tau, x-e) - f(t,x))\rho_\varepsilon(\tau,e)ded\tau \\ &\le [f]_{\mathcal{C}^{1/2,1}}\varepsilon \le M\varepsilon. \end{aligned}$$

This together with (4.14) gives that in $\bar{Q}_{T-\Delta}$,

$$\max\{-\partial_t u_\varepsilon + g(t,x,D_x u_\varepsilon, D_x^2 u_\varepsilon), u_\varepsilon - f(t,x)\} \le M\varepsilon.$$

We then substitute u_ε into the consistency error estimate (4.6) and use (4.13) to obtain

$$\begin{aligned} &\bar{S}(\Delta, t, x, u_\varepsilon(t,x), u_\varepsilon(t+\Delta, \cdot)) \\ &\le M\varepsilon + C\Delta\left(|\partial_t^2 u_\varepsilon|_0 + |D_x^4 u_\varepsilon|_0 + |\partial_t D_x^2 u_\varepsilon|_0 + \mathcal{R}(u_\varepsilon)\right) \\ &\le C(\varepsilon + \Delta\varepsilon^{-3}), \end{aligned}$$

for $(t,x) \in \bar{Q}_{T-\Delta}$. The comparison result in Proposition 4.1.4 then implies

$$u_\varepsilon - u^\Delta \le \sup_{\bar{Q}_T \setminus \bar{Q}_{T-\Delta}} (u_\varepsilon - u^\Delta)^+ + C(T-t+1)(\varepsilon + \Delta\varepsilon^{-3}) \text{ in } \bar{Q}_T.$$

Next, using estimates (4.11) and (4.12), we obtain that $|u - u_\varepsilon| \le C\varepsilon$ and, thus,

$$\begin{aligned} u - u^\Delta &= (u - u_\varepsilon) + (u_\varepsilon - u^\Delta) \\ &\le C\varepsilon + \sup_{\bar{Q}_T \setminus \bar{Q}_{T-\Delta}} (u_\varepsilon - u^\Delta)^+ + C(\varepsilon + \Delta\varepsilon^{-3}) \\ &\le \sup_{\bar{Q}_T \setminus \bar{Q}_{T-\Delta}} (u - u^\Delta)^+ + C(\varepsilon + \Delta\varepsilon^{-3}) \text{ in } \bar{Q}_T. \end{aligned}$$

By choosing $\varepsilon = \Delta^{\frac{1}{4}}$, we further obtain

$$u - u^{\Delta} \leq \sup_{\bar{Q}_T \setminus \bar{Q}_{T-\Delta}} (u - u^{\Delta})^+ + C\Delta^{\frac{1}{4}} \leq C\Delta^{\frac{1}{4}} \quad \text{in } \bar{Q}_T,$$

where the last inequality follows from the estimate (4.8) in Lemma 4.2.1. ■

4.2.2 Lower bound for the approximation error

To obtain a lower bound of $u - u^{\Delta}$, we apply again the *optimal switching approximation* method. However, different from the case in Section 3.3.2 where the standard optimal switching systems are used to approximate an HJB equation, the switching systems we will use in this chapter consist of an extra obstacle term in order to approximate HJB variational inequalities. We call this type of switching system an *obstacle switching system* and momentarily use its properties in the section while leaving the details of their proofs in Section 4.3.

Follow the same arguments as in Section 3.3.2, we can write the equation (4.1) as

$$\begin{cases} \max\{-\partial_t u + \sup_{q \in K_\infty} \mathcal{L}^q \left(t, x, D_x u, D_x^2 u\right), u - f(t,x)\} = 0 & \text{in } Q_T; \\ u(T,x) = U(x) & \text{in } \mathbb{R}^d, \end{cases} \tag{4.15}$$

where $K_\infty = \{q_1, q_2, q_3, ...\}$ is a countable dense subset of some compact set in $\mathbb{R}^d$.

Next, for $m \in \mathbb{Z}^+$, we consider the following equations to approximate equation (4.15),

$$\begin{cases} \max\{-\partial_t u^m + \sup_{q \in K_m} \mathcal{L}^q \left(t, x, D_x u^m, D_x^2 u^m\right), u^m - f(t,x)\} = 0 & \text{in } Q_T; \\ u^m(T,x) = U(x) & \text{in } \mathbb{R}^d, \end{cases} \tag{4.16}$$

where $K_m := \{q_1, ..., q_m\} \subset K_\infty$ consisting the first m points in K_∞ and satisfying $\cup_{m \geq 1} K_m = K_\infty$. It then follows from standard optimal stopping control problem (see [11, 31]) that (4.16) admits a unique viscosity solution $u^m \in \mathcal{C}_b^1(\bar{Q}_T)$, with $|u^m|_1 \leq C$ independent of m. Then, Arzela-Ascoli's theorem yields that there exists a subsequence of $\{u^m\}$, still denoted as $\{u^m\}$, such that, as $m \to \infty$.

$$u^m \to u \text{ locally uniformly in } \bar{Q}_T. \tag{4.17}$$

Similar to (3.37), we have that for $m \in \mathbb{Z}^+$,

$$|u(t,\cdot) - u^m(t,\cdot)|_0 \leq C\sqrt{T-t}. \tag{4.18}$$

We then construct a sequence of (local) smooth supersolutions to approximate the solution u^m of (4.16). To this end, we consider the following m dimensional obstacle switching system in Q_T:

$$\max\left\{-\partial_t v_i + \mathcal{L}^{q_i}(t, x, D_x v_i, D_x^2 v_i), v_i - f, v_i - \mathcal{M}_i^k v\right\} = 0, \quad i \in \mathcal{I} := \{1, ..., m\} \tag{4.19}$$

with

$$v_i(T,x) = U(x), \quad i \in \mathcal{I} \tag{4.20}$$

where $\mathcal{M}_i^k v := \min_{j\neq i,\ j\in\mathcal{I}}\{v_j + k\}$, for some constant $k>0$ representing the switching cost.

By the results from the next section, we then have the following existence, uniqueness and regularity results for the solution v of obstacle switching system (4.19)-(4.20), and a comparison between v and u^m.

Proposition 4.2.5 *Suppose that Assumption 3.2.1 and 4.1.1 are satisfied. Then, for any $m \in \mathbb{Z}^+$, there exists a unique viscosity solution $v = (v_1, \dots, v_m)$ of (4.19)-(4.20), with $|v|_1 \le C$. Moreover, for k small enough,*

$$0 \le v_i - u^m \le Ck^{\frac{1}{3}} \quad \textit{in } \bar{Q}_T, \quad i \in \mathcal{I}. \tag{4.21}$$

Proof. The existence, uniqueness and regularity result of the viscosity solution v is given by Theorem 4.3.5 in the next section. To get the estimates (4.21), we first check that $w = (u^m, ..., u^m)$ is a subsolution of (4.19), then comparison result for the obstacle switching system (which is implied by Theorem 4.3.4) yields $u^m \le v_i$ for $i \in \mathcal{I}$.

To derive the other bound, we follow the same regularization procedure as in Section 4.2.1. For small enough $\varepsilon > 0$, we consider the following perturbed system of (4.19) in $Q_{T+\varepsilon^2}$ for $i \in \mathcal{I}$:

$$\max\left\{-\partial_t v_i^\varepsilon + \sup_{(\tau,e)\in\Theta^\varepsilon} \mathcal{L}^{q_i}(t+\tau, x+e, D_x v_i^\varepsilon, D_x^2 v_i^\varepsilon), v_i^\varepsilon - f, v_i^\varepsilon - \mathcal{M}_i^k v^\varepsilon\right\} = 0, \tag{4.22}$$

with

$$v_i^\varepsilon(T+\varepsilon^2, x) = U(x), \tag{4.23}$$

where $\Theta^\varepsilon = [-\varepsilon^2, 0] \times \varepsilon B(0,1)$. Note that here we extend the coefficients σ, b, f and L appropriately. It then follows from Theorem 4.3.5 and Theorem 4.3.4 that (4.22)-(4.23) admits a unique viscosity solution $v^\varepsilon = (v_1^\varepsilon, \dots, v_m^\varepsilon)$ such that for $i \in \mathcal{I}$,

$$|v_i^\varepsilon|_1 \le C \text{ and } |v_i^\varepsilon - v_i| \le C\varepsilon \text{ in } \bar{Q}_T. \tag{4.24}$$

It then follows from (4.22) that for any $(\tau, e) \in \Theta^\varepsilon$,

$$-\partial_t v_i^\varepsilon + \mathcal{L}^{q_i}(t+\tau, x+e, D_x v_i^\varepsilon, D_x^2 v_i^\varepsilon) \le 0 \quad \text{in} \quad Q_{T+\varepsilon^2}, \quad i \in \mathcal{I}.$$

For each $i \in \mathcal{I}$, define $v_{i,\varepsilon} := v_i^\varepsilon * \rho_\varepsilon$, where ρ_ε is defined in (3.30). By the same argument as in Section 4.2.1, we have

$$-\partial_t v_{i,\varepsilon} + \mathcal{L}^{q_i}(t, x, D_x v_{i,\varepsilon}, D_x^2 v_{i,\varepsilon}) \le 0 \quad \text{in} \quad Q_T, \quad i \in \mathcal{I}. \tag{4.25}$$

On the other hand, it follows again from (4.22) that for any $i \in \mathcal{I}$, $v_i^\varepsilon \le \mathcal{M}_i^k v^\varepsilon = \min_{j\neq i, j\in\mathcal{I}} v_j^\varepsilon + k$ in $Q_{T+\varepsilon^2}$, which implies that

$$|v_i^\varepsilon - v_j^\varepsilon|_0 \le k, \quad i,j \in \mathcal{I}.$$

Then by standard properties of mollifiers, we have

$$|\partial_t v_{i,\varepsilon} - \partial_t v_{j,\varepsilon}|_0 \le Ck\varepsilon^{-2}; \quad |D_x^n v_{i,\varepsilon} - D_x^n v_{j,\varepsilon}|_0 \le Ck\varepsilon^{-n}, \quad n \in \mathbb{N}, \quad i,j \in \mathcal{I},$$

These estimates then yield that

$$|-\partial_t v_{j,\varepsilon}+\mathcal{L}^{q_j}(t,x,D_x v_{j,\varepsilon},D_x^2 v_{j,\varepsilon})+\partial_t v_{i,\varepsilon}-\mathcal{L}^{q_j}(t,x,D_x v_{i,\varepsilon},D_x^2 v_{i,\varepsilon})|\leq Ck(\varepsilon^{-2}+\varepsilon^{-1})$$

This together with (4.25) gives that

$$-\partial_t v_{i,\varepsilon}+\mathcal{L}^{q_j}(t,x,D_x v_{i,\varepsilon},D_x^2 v_{i,\varepsilon})\leq Ck(\varepsilon^{-2}+\varepsilon^{-1})\leq Ck\varepsilon^{-2}\quad \text{in } Q_T,\quad i,j\in\mathcal{I},$$

which means

$$-\partial_t v_{i,\varepsilon}+\sup_{q\in K_m}\mathcal{L}^{q}(t,x,D_x v_{i,\varepsilon},D_x^2 v_{i,\varepsilon})\leq Ck\varepsilon^{-2}\quad \text{in } Q_T,\quad i\in\mathcal{I}.$$

Next, since for any $i\in\mathcal{I}$, $v_i^\varepsilon\leq f$ in $Q_{T+\varepsilon^2}$, we obtain that for $(t,x)\in Q_T$,

$$v_{i,\varepsilon}(t,x)-f(t,x)\leq\int_{-\varepsilon^2<\tau<0}\int_{|e|<\varepsilon}(f(t-\tau,x-e)-f(t,x))\rho_\varepsilon(\tau,e)ded\tau\leq M\varepsilon.$$

From the above two inequalities we can see that for any $i\in\mathcal{I}$, $v_{i,\varepsilon}-(T-t)Ck\varepsilon^{-2}-M\varepsilon$ is a subsolution of (4.16) in Q_T with terminal value $v_{i,\varepsilon}(T,\cdot)-M\varepsilon$. Then by standard continuous dependence result for (4.16), we have for $i\in\mathcal{I}$,

$$\begin{aligned}v_{i,\varepsilon}-u^m&\leq|(v_{i,\varepsilon}(T,\cdot)-M\varepsilon-u^m(T,\cdot))^+|_0+(T-t)Ck\varepsilon^{-2}+M\varepsilon\\&\leq|v_{i,\varepsilon}(T,\cdot)-v_i^\varepsilon(T+\varepsilon^2,\cdot)|_0+C(\varepsilon+k\varepsilon^{-2})\quad\text{in } Q_T.\end{aligned}$$

Hence by the properties of mollifiers and regularity of v^ε, we have for $i\in\mathcal{I}$,

$$\begin{aligned}v_i^\varepsilon-u^m&=v_i^\varepsilon-v_{i,\varepsilon}+v_{i,\varepsilon}-u^m\\&\leq C\varepsilon+|v_{i,\varepsilon}(T,\cdot)-v_i^\varepsilon(T,\cdot)|_0\\&\quad+|v_i^\varepsilon(T,\cdot)-v_i^\varepsilon(T+\varepsilon^2,\cdot)|_0+C(\varepsilon+k\varepsilon^{-2})\\&\leq C(\varepsilon+k\varepsilon^{-2})\quad\text{in } Q_T.\end{aligned}$$

We then choose $\varepsilon=k^{\frac{1}{3}}$ and finish the proof. ■

Proposition 4.2.6 *Suppose that Assumption 3.2.1 and 4.1.1 are satisfied. Let $m\in\mathbb{Z}^+$, $u^\Delta\in\mathcal{C}_b(\bar{Q}_T)$ be the unique solution of the approximation scheme (4.3) and $u^m\in\mathcal{C}_b^1(\bar{Q}_T)$ be the unique viscosity solution of equation (4.16). Then,*

$$u^\Delta-u^m\leq\sup_{\bar{Q}_T\setminus\bar{Q}_{T-\Delta}}(u^\Delta-u^m)^++C\Delta^{\frac{1}{10}}\ \text{ in } \bar{Q}_T.$$

Proof. In analogy to (4.22) but in the opposite direction, for small enough $\varepsilon>0$, we perturb the coefficients of the system (4.19)-(4.20), and consider the following obstacle switching system in $Q_{T+\varepsilon^2}$ for $i\in\mathcal{I}$:

$$\max\left\{-\partial_t v_i^\varepsilon+\inf_{(\tau,e)\in\Theta^\varepsilon}\mathcal{L}^{q_i}(t+\tau,x+e,D_x v_i^\varepsilon,D_x^2 v_i^\varepsilon),v_i^\varepsilon-f,v_i^\varepsilon-\mathcal{M}_i^k v^\varepsilon\right\}=0,\tag{4.26}$$

with

$$v_i^\varepsilon(T+\varepsilon^2,x)=U(x).\tag{4.27}$$

It then follows again from Theorem 4.3.5 and Theorem 4.3.4 that (4.26)-(4.27) admits a unique viscosity solution $v^\varepsilon = (v_1^\varepsilon, \ldots, v_m^\varepsilon)$, with $|v_i^\varepsilon|_1 \le C$ and, moreover, for each $i \in \mathcal{I}$,

$$|v_i^\varepsilon - v_i| \le C\varepsilon \text{ in } \bar{Q}_T.$$

In turn, this and inequality (4.21) imply that, for each $i \in \mathcal{I}$,

$$|v_i^\varepsilon - u^m| \le |v_i^\varepsilon - v_i| + |v_i - u^m| \le C(\varepsilon + k^{\frac{1}{3}}) \text{ in } \bar{Q}_T. \tag{4.28}$$

Next, we regularize v^ε and define $v_{i,\varepsilon} := v_i^\varepsilon * \rho_\varepsilon$ in $\bar{Q}_T$ for $i \in \mathcal{I}$. Then, $v_{i,\varepsilon} \in \mathcal{C}_b^\infty(\bar{Q}_T)$,

$$|v_{i,\varepsilon} - v_i^\varepsilon|_0 \le C\varepsilon, \tag{4.29}$$

and moreover, for positive integer m and multi-index n,

$$|\partial_t^m D_x^n v_{i,\varepsilon}|_0 \le C\varepsilon^{1-2m-|n|}. \tag{4.30}$$

Define $w_\varepsilon := \min_{i\in\mathcal{I}} v_{i,\varepsilon}$ in $\bar{Q}_T$, which is smooth in $\bar{Q}_T$ except for finitely many points. Then, (4.28) and (4.29) yield

$$|u^m - w_\varepsilon| \le C(\varepsilon + k^{\frac{1}{3}}) \text{ in } \bar{Q}_T. \tag{4.31}$$

Now we fix any $(t,x) \in \bar{Q}_{T-\Delta}$ and try to obtain a lower bound for $\bar{S}(\Delta, t, x, w_\varepsilon(t,x), w_\varepsilon(t+\Delta,\cdot))$. For this, we consider two cases when $w_\varepsilon(t,x) \ge f(t,x) - k$ and $w_\varepsilon(t,x) < f(t,x) - k$,

(i) If $w_\varepsilon(t,x) \ge f(t,x) - k$, it is obvious that $\bar{S}(\Delta, t, x, w_\varepsilon(t,x), w_\varepsilon(t + \Delta,\cdot)) \ge w_\varepsilon(t,x) - f(t,x) \ge -k$.

(ii) If $w_\varepsilon(t,x) < f(t,x) - k$, let $j = \arg\min_{i\in\mathcal{I}} v_{i,\varepsilon}(t,x)$. Then we obtain that

$$v_{j,\varepsilon}(t,x) = w_\varepsilon(t,x) < f(t,x) - k,$$

and

$$v_{j,\varepsilon}(t,x) - \mathcal{M}_j^k v_\varepsilon(t,x) = \max_{i\ne j, i\in\mathcal{I}} \{v_{j,\varepsilon}(t,x) - v_{i,\varepsilon}(t,x) - k\} \le -k.$$

In turn, inequality (4.29) implies that

$$v_j^\varepsilon(t,x) - f(t,x) \le v_{j,\varepsilon}(t,x) - f(t,x) + C\varepsilon < -k + C\varepsilon,$$

and

$$v_j^\varepsilon(t,x) - \mathcal{M}_j^k v^\varepsilon(t,x) \le v_{j,\varepsilon}(t,x) - \mathcal{M}_j^k v_\varepsilon(t,x) + C\varepsilon \le -k + C\varepsilon.$$

Furthermore, since $|v^\varepsilon|_1, |f|_1 \le C$, we have for any $(\tau, e) \in \Theta^\varepsilon$,

$$\begin{aligned} v_j^\varepsilon(t-\tau, x-e) - f(t-\tau, x-e) &\le v_j^\varepsilon(t,x) - f(t,x) + C(|\tau|^{\frac{1}{2}} + |e|) \\ &< -k + C\varepsilon + 2C\varepsilon, \end{aligned}$$

and similarly,

$$v_j^\varepsilon(t-\tau, x-e) - \mathcal{M}_j^k v^\varepsilon(t-\tau, x-e) \le -k + C\varepsilon + 2C\varepsilon.$$

By choosing $k = 4C\varepsilon$, we then obtain that, for any $(\tau, e) \in \Theta^\varepsilon$,

$$v_j^\varepsilon(t-\tau, x-e) - f(t-\tau, x-e) < 0,$$

and

$$v_j^\varepsilon(t-\tau, x-e) - \mathcal{M}_j^k v^\varepsilon(t-\tau, x-e) < 0.$$

Therefore, the points $(t-\tau, x-e)$, for $(\tau, e) \in \Theta^\varepsilon$, are in the continuation region of (4.26). Then we have that, for $(\tau, e) \in \Theta^\varepsilon$,

$$-\partial_t v_j^\varepsilon(t-\tau, x-e) + \mathcal{L}^{q_j}\left(t, x, D_x v_j^\varepsilon(t-\tau, x-e), D_x^2 v_j^\varepsilon(t-\tau, x-e)\right) \geq 0.$$

Using the definition of $v_{j,\varepsilon}$ and that $\mathcal{L}^{q_j}$ is linear in $D_x v_j^\varepsilon$ and $D_x^2 v_j^\varepsilon$, we further have

$$\begin{aligned}
&-\partial_t v_{j,\varepsilon}(t,x) + \mathcal{L}^{q_j}\left(t, x, D_x v_{j,\varepsilon}(t,x), D_x^2 v_{j,\varepsilon}(t,x)\right) \\
=& \int_{-\varepsilon^2<\tau<0}\int_{|e|<\varepsilon} \left(-\partial_t v_j^\varepsilon(t-\tau, x-e) + \mathcal{L}^{q_j}\left(t, x, D_x v_j^\varepsilon(t-\tau, x-e), D_x^2 v_j^\varepsilon(t-\tau, x-e)\right)\right) \\
&\times \rho_\varepsilon(\tau, e) de d\tau \geq 0. \qquad (4.32)
\end{aligned}$$

Since the definition of j implies that $w_\varepsilon(t,x) = v_{j,\varepsilon}(t,x)$ and $w_\varepsilon(t+\Delta, \cdot) \leq v_{j,\varepsilon}(t+\Delta, \cdot)$, we apply Proposition 4.1.3(ii) (iii) and estimate (4.30), (4.32) to obtain that

$$\begin{aligned}
&\bar{S}(\Delta, t, x, w_\varepsilon(t,x), w_\varepsilon(t+\Delta, \cdot)) \\
\geq\ & \bar{S}(\Delta, t, x, v_{j,\varepsilon}(t,x), v_{j,\varepsilon}(t+\Delta, \cdot)) \\
\geq\ & -\partial_t v_{j,\varepsilon}(t,x) + g(t, x, D_x v_{j,\varepsilon}(t,x), D_x^2 v_{j,\varepsilon}(t,x)) - C\Delta\varepsilon^{-3} \\
=\ & -\partial_t v_{j,\varepsilon}(t,x) + \sup_{q\in\mathbb{R}^d} \mathcal{L}^q(t, x, D_x v_{j,\varepsilon}(t,x), D_x^2 v_{j,\varepsilon}(t,x)) - C\Delta\varepsilon^{-3} \\
\geq\ & -\partial_t v_{j,\varepsilon}(t,x) + \mathcal{L}^{q_j}(t, x, D_x v_{j,\varepsilon}(t,x), D_x^2 v_{j,\varepsilon}(t,x)) - C\Delta\varepsilon^{-3} \\
\geq\ & -C\Delta\varepsilon^{-3}.
\end{aligned}$$

Thus, by choosing $k = 4C\varepsilon$ and combining case (i) and (ii), we have

$$\bar{S}(\Delta, t, x, w_\varepsilon(t,x), w_\varepsilon(t+\Delta, \cdot)) \geq -C(\Delta\varepsilon^{-3} + \varepsilon).$$

Note that the right hand side of the above inequality does not depend on (t,x), thus this inequality holds in $\bar{Q}_{T-\Delta}$. In turn, the comparison result in Proposition 4.1.4 implies that

$$u^\Delta - w_\varepsilon \leq \sup_{\bar{Q}_T \setminus \bar{Q}_{T-\Delta}} (u^\Delta - w_\varepsilon)^+ + C(T-t+1)(\Delta\varepsilon^{-3} + \varepsilon) \text{ in } \bar{Q}_T.$$

Combining the above inequality with (4.31), we further get

$$\begin{aligned}
u^{\Delta} - u^m &= (u^{\Delta} - w_\varepsilon) + (w_\varepsilon - u^m) \\
&\le \sup_{\bar{Q}_T \backslash \bar{Q}_{T-\Delta}} (u^{\Delta} - w_\varepsilon)^+ + C(T-t+1)(\Delta\varepsilon^{-3} + \varepsilon) + C(\varepsilon + k^{\frac{1}{3}}) \\
&\le \sup_{\bar{Q}_T \backslash \bar{Q}_{T-\Delta}} (u^{\Delta} - u^m)^+ + C(\varepsilon^{\frac{1}{3}} + \Delta\varepsilon^{-3}) \\
&\le \sup_{\bar{Q}_T \backslash \bar{Q}_{T-\Delta}} (u^{\Delta} - u^m)^+ + C\Delta^{\frac{1}{10}} \text{ in } \bar{Q}_T,
\end{aligned}$$

where we used $k = 4C\varepsilon$ in the second to last inequality, and choose $\varepsilon = \Delta^{\frac{3}{10}}$ in the last inequality. ■

We are now ready to obtain the lower bound for the approximation error.

Theorem 4.2.7 *Suppose that Assumption 3.2.1 and 4.1.1 are satisfied. Let $u^{\Delta} \in \mathcal{C}_b(\bar{Q}_T)$ be the unique solution of the approximation scheme (4.3) and $u \in \mathcal{C}_b^1(\bar{Q}_T)$ be the unique viscosity solution of equation (4.1). Then,*

$$u - u^{\Delta} \ge -C\Delta^{\frac{1}{10}} \textit{ in } \bar{Q}_T.$$

Proof. Applying Proposition 4.2.6 to the sequence $\{u^m\}$, we get for $m \in \mathbb{Z}^+$,

$$\begin{aligned}
u^{\Delta} - u &= (u^{\Delta} - u^m) + (u^m - u) \\
&\le \sup_{\bar{Q}_T \backslash \bar{Q}_{T-\Delta}} (u^{\Delta} - u^m)^+ + C\Delta^{\frac{1}{10}} + (u^m - u) \\
&\le \sup_{\bar{Q}_T \backslash \bar{Q}_{T-\Delta}} (u^{\Delta} - u)^+ + \sup_{\bar{Q}_T \backslash \bar{Q}_{T-\Delta}} (u - u^m)^+ + C\Delta^{\frac{1}{10}} + u^m - u \\
&\le C\Delta^{\frac{1}{10}} + u^m - u \quad \text{in } \bar{Q}_T,
\end{aligned}$$

where we used estimate (4.8) and (4.18) in the last inequality. Sending $m \to \infty$ and using (4.17), we conclude. ■

4.3 Well-posedness, Regularity, and Continuous Dependence for an Obstacle Switching System

In this section we construct the well-posedness, regularity, and continuous dependence results for general obstacle switching systems, which include equation (4.19), (4.22) and (4.26) as special cases. These results are vitally important in the previous section when obtaining the lower bound of approximation error.

For $T > 0$ and $m \in \mathbb{Z}^+$, we consider the following m dimensional general obstacle switching system in Q_T:

$$\max\left\{-\partial_t u_i + \sup_{\alpha\in\mathcal{A}} \inf_{\beta\in\mathcal{B}} \mathcal{L}_i^{\alpha,\beta}(t, x, D_x u_i, D_x^2 u_i), u_i - f, u_i - \mathcal{M}_i^k u\right\} = 0, \quad i \in \mathcal{I}, \tag{4.33}$$

with terminal condition

$$u_i(T,x) = U(x), \quad i \in \mathcal{I}, \tag{4.34}$$

where

$$\mathcal{L}_i^{\alpha,\beta}(t,x,p,X) = -\frac{1}{2}\mathrm{tr}\left(\sigma_i^{\alpha,\beta}\sigma_i^{\alpha,\beta^T}(t,x)X\right) - b_i^{\alpha,\beta}(t,x)\cdot p - L_i^{\alpha,\beta}(t,x),$$

$\mathcal{A}$ and $\mathcal{B}$ are compact metric spaces. Note that this obstacle switching system is a general version of (4.19), (4.22), and (4.26). We then make the following assumption:

Assumption 4.3.1 *There exists a constant $\bar{C} > 0$ independent of α, β, i and t such that for any α, β, i, and t,*

$$|\sigma_i^{\alpha,\beta}(t,\cdot)|_1, |b_i^{\alpha,\beta}(t,\cdot)|_1, |L_i^{\alpha,\beta}(t,\cdot)|_1, |f|_1, |U|_1 \le \bar{C}.$$

We firstly give the existence of the solution of (4.33)-(4.34), then proceed to the continuous dependence on the coefficients, which implies the standard comparison principle and uniqueness, and finally obtain the regularity for the solution based on continuous dependence results.

Proposition 4.3.2 *Suppose that Assumption 4.3.1 is satisfied. Then there exists a bounded viscosity solution u of (4.33)-(4.34), with $|u|_0 \le C$ depending only on T and $\bar{C}$.*

Proof. We fist claim that the viscosity solution of (4.33)-(4.34) can be represented as the value function of a mixed optimal stopping and switching problem, and show that the value function is uniformly bounded. We then give a sketch of proof that the value function is indeed a viscosity solution of (4.33)-(4.34). Specifically, the viscosity solution of (4.33)-(4.34) can be represented on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\ge 0}, \mathbb{P})$ by

$$u_i(t,x) = \inf_{\substack{\theta\in\Theta_i[t,T] \\ \tau\in\mathcal{T}[t,T] \\ \alpha\in\mathcal{A}[t,T]}} \sup_{\beta\in\mathcal{B}[t,T]} J(t,x;\theta,\tau,\alpha,\beta), \tag{4.35}$$

with the cost function

$$\begin{aligned} J(t,x;\theta,\tau,\alpha,\beta) = \mathbb{E}^{t,x}\Bigg[&\int_t^\tau L_{\theta_s}^{\alpha_s,\beta_s}\left(s, X_s^{\alpha,\beta;\theta}\right) ds + k\sum_{n\ge 1}\mathbf{1}_{\{\tau_n\le\tau\}} \\ &+ f(\tau, X_\tau^{\alpha,\beta;\theta})\mathbf{1}_{\{\tau<T\}} + U\left(X_T^{\alpha,\beta;\theta}\right)\mathbf{1}_{\{\tau=T\}}|\mathcal{F}_t\Bigg], \end{aligned} \tag{4.36}$$

and the diffusion process X driven by

$$dX_s^{\alpha,\beta;\theta} = b_{\theta_s}^{\alpha_s,\beta_s}(s, X_s^{\alpha,\beta;\theta})ds + \sigma_{\theta_s}^{\alpha_s,\beta_s}(s, X_s^{\alpha,\beta;\theta})dW_s,$$

where $\mathcal{A}[t,T]$ and $\mathcal{B}[t,T]$ are $\mathcal{A}$ and $\mathcal{B}$-valued progressive-measurable processes respectively, $\mathcal{T}[t,T]$ is the set of stopping times valued in $[t,T]$, and $\Theta_i[t,T]$ is the space of all admissible continuous switching control processes on $[t,T]$ starting from i. Specifically, for any admissible switching control process, there is a pair of corresponding sequence $\{\xi_n, \tau_n\}_{n\geq 0}$ such that $\{\tau_n\}_{n\geq 0}$ is a sequence of nondecreasing stopping time with $\tau_0 = t$ a.s. and that each ξ_n is a $\mathcal{F}_{\tau_n}$-measurable random variable valued in $\mathcal{I}$, with $\xi_0 = i$ and $\xi_n \neq \xi_{n+1}$ a.s. Then an admissible switching control process θ in $\Theta_i[t,T]$ is identified as

$$\theta_t = \Sigma_{n\geq 1}\xi_{n-1}\mathbf{1}_{[\tau_{n-1},\tau_n)}(t).$$

The representation given by (4.35)-(4.36) immediately shows that $u \geq -C$ since 4.3.1 holds, for some constant C depending only on T and $\bar{C}$. That $u \leq C$ can be obtained by simply choosing $\tau = t$, i.e. stop immediately, to obtain that $u \leq f$.

We remain to prove that the value function u is indeed a viscosity solution of (4.33)-(4.34). The proof is an extension of the arguments in [64] and thus we only give a sketch of proof. The terminal condition (4.34) is immediately satisfied by letting $t = T$ in (4.35)-(4.36). We now prove that u is a viscosity solution of (4.33) in Q_T. To this end, we appeal to the dynamic programming principle, which takes the form of the following two inequalities:

$$\begin{aligned} u_i(t,x) \leq & \inf_{\substack{\theta\in\Theta_i[t,T] \\ \tau\in\mathcal{T}[t,T] \\ \alpha\in\mathcal{A}[t,T]}} \sup_{\beta\in\mathcal{B}[t,T]} \mathbb{E}^{t,x}\left[\int_t^{\tau\wedge\nu} L_{\theta_s}^{\alpha_s,\beta_s}\left(s, X_s^{\alpha,\beta;\theta}\right)ds + k\sum_{n\geq 1}\mathbf{1}_{\{\tau_n\leq\tau\wedge\nu\}}\right. \\ & + \left. f(\tau, X_\tau^{\alpha,\beta;\theta})\mathbf{1}_{\{\tau<\nu\}} + u_{\theta_\nu}^*\left(\nu, X_\nu^{\alpha,\beta;\theta}\right)\mathbf{1}_{\{\nu\leq\tau\}}|\mathcal{F}_t\right], \end{aligned} \tag{4.37}$$

$$\begin{aligned} u_i(t,x) \geq & \inf_{\substack{\theta\in\Theta_i[t,T] \\ \tau\in\mathcal{T}[t,T] \\ \alpha\in\mathcal{A}[t,T]}} \sup_{\beta\in\mathcal{B}[t,T]} \mathbb{E}^{t,x}\left[\int_t^{\tau\wedge\nu} L_{\theta_s}^{\alpha_s,\beta_s}\left(s, X_s^{\alpha,\beta;\theta}\right)ds + k\sum_{n\geq 1}\mathbf{1}_{\{\tau_n\leq\tau\wedge\nu\}}\right. \\ & + \left. f(\tau, X_\tau^{\alpha,\beta;\theta})\mathbf{1}_{\{\tau<\nu\}} + u_{\theta_\nu,*}\left(\nu, X_\nu^{\alpha,\beta;\theta}\right)\mathbf{1}_{\{\nu\leq\tau\}}|\mathcal{F}_t\right], \end{aligned} \tag{4.38}$$

where $\nu \in \mathcal{T}[t,T]$ is any stopping time valued in $[t,T]$, and u_i^* and $u_{i,*}$ are upper- and lower-semicontinuous envelope of u_i respectively. We then divide the proof into two steps.

1. We first show that u_i is a viscosity subsolution. We have shown that $u_i \leq f$, which implies that $u_i^* - f \leq 0$. By taking the immediate switching control $\tau_1 = t$, $\xi_1 = j \neq i$, $\tau_n = \infty$ for $n \geq 2$, and choosing $\nu = t$ in (4.37), we obtain $u_i \leq k + u_j^*$, which implies that $u_i^* \leq k + u_j^*$ and thus $u_i^* - \mathcal{M}_i^k u^* \leq 0$. Next, let $(t_0, x_0) \in Q_T$ and $\varphi \in \mathcal{C}^{1,2}(Q_T)$ be such that

$$0 = (u_i^* - \varphi)(t_0, x_0) = \max_{Q_T}(u_i^* - \varphi).$$

Consider a sequence $(t_n, x_n)_{n\geq 1} \in Q_T$ such that

$$(t_n, x_n) \to (t_0, x_0) \text{ and } u_i(t_n, x_n) \to u_i^*(t_0, x_0).$$

Denote $\eta_n := (\varphi - u_i)(t_n, x_n)$ and we have $\eta_n \downarrow 0$. Let $h_n := \sqrt{\eta_n}\mathbf{1}_{\{\eta_n\neq 0\}} + n^{-1}\mathbf{1}_{\{\eta_n=0\}}$. Then for large enough n, $t_n + h_n < T$ and thus $\nu_n := t_n + h_n \in$

$\mathcal{T}[t,T]$. We then take $\tau = T$, and $\bar{\theta}_s \equiv i$, i.e. $\tau_n = \infty$ for $n \geq 1$ in (4.37) to get

$$\begin{aligned}
-\eta_n + \varphi(t_n, x_n) \leq & \inf_{\alpha \in \mathcal{A}[t,T]} \sup_{\beta \in \mathcal{B}[t,T]} \mathbb{E}^{t_n, x_n} \left[\int_{t_n}^{t_n + h_n} L_i^{\alpha_s, \beta_s} \left(s, X_s^{\alpha, \beta; \bar{\theta}} \right) ds \right. \\
& + \left. u_i^* \left(t_n + h_n, X_{t_n + h_n}^{\alpha, \beta; \bar{\theta}} \right) | \mathcal{F}_{t_n} \right] \\
\leq & \inf_{\alpha \in \mathcal{A}[t,T]} \sup_{\beta \in \mathcal{B}[t,T]} \mathbb{E}^{t_n, x_n} \left[\int_{t_n}^{t_n + h_n} L_i^{\alpha_s, \beta_s} \left(s, X_s^{\alpha, \beta; \bar{\theta}} \right) ds \right. \\
& + \left. \varphi \left(t_n + h_n, X_{t_n + h_n}^{\alpha, \beta; \bar{\theta}} \right) | \mathcal{F}_{t_n} \right].
\end{aligned}$$

By applying Itô's lemma to φ, dividing h_n in both side and letting $n \to \infty$, we obtain that

$$-\partial_t \varphi(t_0, x_0) + \sup_{\alpha \in \mathcal{A}} \inf_{\beta \in \mathcal{B}} \mathcal{L}_i^{\alpha, \beta}(t_0, x_0, D_x \varphi(t_0, x_0), D_x^2 \varphi(t_0, x_0)) \leq 0.$$

2. We next prove that u_i is a viscosity supersolution by contradiction. Let $(t_0, x_0) \in Q_T$ and $\varphi \in \mathcal{C}^{1,2}(Q_T)$ be such that

$$0 = (u_{i,*} - \varphi)(t_0, x_0) = \text{strict } \min_{Q_T} (u_{i,*} - \varphi).$$

We assume to the contrary that

$$(\varphi - f)(t_0, x_0) < 0, \quad (\varphi - \mathcal{M}_i^k u_*)(t_0, x_0) < 0,$$

and

$$-\partial_t \varphi(t_0, x_0) + \sup_{\alpha \in \mathcal{A}} \inf_{\beta \in \mathcal{B}} \mathcal{L}_i^{\alpha, \beta}(t_0, x_0, D_x \varphi(t_0, x_0), D_x^2 \varphi(t_0, x_0)) < 0.$$

Then there exist constants $h, \delta > 0$ such that on $\mathcal{N}_h := [t_0, t_0 + h] \times hB(x_0)$,

$$\varphi \leq f - \delta, \quad \varphi \leq \mathcal{M}_i^k u_* - \delta,$$

and for any $\alpha \in \mathcal{A}$ there exists a $\beta(\alpha) \in \mathcal{B}$ such that

$$-\partial_t \varphi + \mathcal{L}_i^{\alpha, \beta(\alpha)}[\varphi] := -\partial_t \varphi + \mathcal{L}_i^{\alpha, \beta(\alpha)}(t, x, D_x \varphi, D_x^2 \varphi) \leq 0,$$

where $t_0 + h < T$ and $B(x_0)$ is the d-dimensional unit ball centered at x_0. Moreover, since (t_0, x_0) is a strict minimizer, we denote

$$\gamma := \min_{\partial \mathcal{N}_h} (u_{i,*} - \varphi) > 0.$$

Consider a sequence $(t_n, x_n)_{n \geq 1} \in Q_T$ such that

$$(t_n, x_n) \to (t_0, x_0) \text{ and } u_i(t_n, x_n) \to u_{i,*}(t_0, x_0).$$

Denote $\eta_n := (u_i - \varphi)(t_n, x_n)$ and we have $\eta_n \downarrow 0$. We next fix any $\theta \in \Theta_i[t,T], \tau \in \mathcal{T}[t,T]$ and $\alpha \in \mathcal{A}[t,T]$ and define the stopping times

$$\nu_n := \inf\{t > t_n : (t, X_t^{\alpha, \beta(\alpha); \theta}) \notin \mathcal{N}_h\}.$$

By applying Itô's lemma again to φ, we have

$$
\begin{aligned}
&u_i(t_n, x_n)\\
&= \eta_n + \varphi(t_n, x_n)\\
&= \eta_n + \mathbb{E}^{t_n,x_n}\left[\int_{t_n}^{\tau\wedge\tau_1\wedge\nu_n} L_{\theta_s}^{\alpha_s,\beta_s(\alpha)}\left(s, X_s^{\alpha,\beta(\alpha);\theta}\right) ds + \varphi\left(\tau\wedge\tau_1\wedge\nu_n, X_{\tau\wedge\tau_1\wedge\nu_n}^{\alpha,\beta(\alpha);\theta}\right)\right.\\
&+ \left.\int_{t_n}^{\tau\wedge\tau_1\wedge\nu_n}(-\partial_t\varphi + \mathcal{L}_{\theta_s}^{\alpha_s,\beta_s(\alpha)}[\varphi])\left(s, X_s^{\alpha,\beta(\alpha);\theta}\right) ds|\mathcal{F}_{t_n}\right]\\
&\le \eta_n - \delta\wedge\gamma + \mathbb{E}^{t_n,x_n}\left[\int_{t_n}^{\tau\wedge\tau_1\wedge\nu_n} L_{\theta_s}^{\alpha_s,\beta_s(\alpha)}\left(s, X_s^{\alpha,\beta(\alpha);\theta}\right) ds + k\sum_{n\ge1}\mathbf{1}_{\{\tau_n\le\tau\wedge\tau_1\wedge\nu_n\}}\right.\\
&+ \left. f(\tau, X_\tau^{\alpha,\beta(\alpha);\theta})\mathbf{1}_{\{\tau<\tau_1\wedge\nu_n\}} + u_{\theta_{\tau_1\wedge\nu_n},*}\left(\tau_1\wedge\nu_n, X_{\tau_1\wedge\nu_n}^{\alpha,\beta(\alpha);\theta}\right)\mathbf{1}_{\{\tau_1\wedge\nu_n\le\tau\}}|\mathcal{F}_{t_n}\right],
\end{aligned}
$$

where we used that $\varphi \le f-\delta$, $\varphi \le \mathcal{M}_i^k u_* - \delta$, and $-\partial_t\varphi + \mathcal{L}_i^{\alpha,\beta(\alpha)}[\varphi] \le 0$ on $\mathcal{N}_h$. For large enough n, we have $\eta_n - \delta\wedge\gamma < 0$, and since $\theta \in \Theta_i[t,T], \tau \in \mathcal{T}[t,T]$ and $\alpha \in \mathcal{A}[t,T]$ are arbitrary, this provides a contradiction of (4.38) with $\nu = \tau_1 \wedge \nu_n$. ■

Remark 4.3.3 *Besides optimal stopping and switching, problem (4.35) also considers optimal control (α) and parameter uncertainty (β). This makes the obstacle switching system (4.33) very general and robust. On the contrary, the switching system considered in [23] does not include any control or parameter uncertainty, but consists of some nonlocal terms related to discontinuous jumping processes.*

We now give the continuous dependence result for the general obstacle switching system (4.33).

Theorem 4.3.4 *For any $s \in (0,T]$, let $u \in USC(\bar{Q}_s)$ be a bounded from above viscosity subsolution of (4.33) with coefficients $\{\sigma_i^{\alpha,\beta}, b_i^{\alpha,\beta}, L_i^{\alpha,\beta}, f\}$, and $\bar{u} \in LSC(\bar{Q}_s)$ be a bounded from below viscosity supersolution of (4.33) with coefficients $\{\bar{\sigma}_i^{\alpha,\beta}, \bar{b}_i^{\alpha,\beta}, \bar{L}_i^{\alpha,\beta}, \bar{f}\}$. Suppose that Assumption 4.3.1 holds for both sets of coefficients, and that $[u(s,\cdot)]_1 \le M$ or $[\bar{u}(s,\cdot)]_1 \le M$ for some constant M, then there exists a constant C depending only on M, $\bar{C}$, and s such that for each $i \in \mathcal{I}$,*

$$
\begin{aligned}
u_i - \bar{u}_i \le C &\left(\sup_i |(u_i-\bar{u}_i)^+(s,\cdot)|_0 + \sup_{i,\alpha,\beta}\left\{|\sigma-\bar{\sigma}|_0 + |b-\bar{b}|_0\right\}\right.\\
&\left.+ \sup_{i,\alpha,\beta}|L-\bar{L}|_0 + |f-\bar{f}|_0\right) \text{ in } \bar{Q}_s. \qquad (4.39)
\end{aligned}
$$

Proof. The proof is mainly based on the proofs of Theorem A.1 in [39] and Theorem A.3 in [4], and can be regarded as a combination of these two proofs. We give the details here for reader's convenience.

Fix $0 < s \le T$ and in $\bar{Q}_s$, we define functions $v(t,x) := e^t u(t,x)$, $\bar{v}(t,x) := e^t\bar{u}(t,x)$ and $g'(t,x) := e^t g(t,x)$ for $g = L, \bar{L}, f, \bar{f}$. It is then follows that $v \in USC(\bar{Q}_s)$ and $\bar{v} \in LSC(\bar{Q}_s)$ are bounded above viscosity subsolution and bounded below supersolution of the following system with coefficients $\{\sigma, b, L', f'\}$ and $\{\bar{\sigma}, \bar{b}, \bar{L}', \bar{f}'\}$ respectively:

$$\max\left\{-\partial_t v_i + v_i + \sup_{\alpha\in\mathcal{A}}\inf_{\beta\in\mathcal{B}} \mathcal{L}'^{\alpha,\beta}_i(t,x,D_x v_i, D_x^2 v_i), v_i - f', v_i - \mathcal{M}_i^k v\right\} = 0,\ \ i\in\mathcal{I}, \tag{4.40}$$

We now use a doubling variables argument to derive a upper bound for $v_i - \bar{v}_i$ and then will derive for $u_i - \bar{u}_i$ by using back-substitution.

To continue, we define in $[0,s]\times\mathbb{R}^d\times\mathbb{R}^d$ that $\phi(t,x,y) := e^{\lambda(s-t)}\delta|x-y|^2 + \varepsilon(|x|^2+|y|^2)$ and for any $i\in\mathcal{I}$, $\psi^i(t,x,y) := v_i(t,x) - \bar{v}_i(t,y) - \phi(t,x,y)$, where $\lambda,\delta,\varepsilon>0$ are positive constants. Then we let $m^s_{\delta,\varepsilon} = \sup_{i,x,y}\psi^i(s,x,y)^+$ and $m_{\lambda,\delta,\varepsilon} = \sup_{i,t,x,y}\psi^i(t,x,y) - m^s_{\delta,\varepsilon}$. Since v_i and $\bar{v}_i$ are bounded from above and below respectively, by classical arguments there exists $i_0\in\mathcal{I}$, $t_0\in[0,s]$ and $x_0,y_0\in\mathbb{R}^d$ depending on λ,δ and ε such that $\psi^{i_0}(t_0,x_0,y_0) = \sup_{i,t,x,y}\psi^i(t,x,y)$. Moreover, by Lemma A.2 of [4], i_0 can be chosen such that $\bar{v}_{i_0}(t_0,y_0) < \mathcal{M}_{i_0}\bar{v}(t_0,y_0)$. Loosely speaking this means that we can now ignore the $\bar{v}_i - \mathcal{M}_i\bar{v}$ parts and proceed as if working with standard variational inequalities. Note that by letting $y=x$, we have for each i,

$$m^s_{\delta,\varepsilon} + m_{\lambda,\delta,\varepsilon} = \psi^{i_0}(t_0,x_0,y_0) \geq v_i(t,x) - \bar{v}_i(t,x) - 2\varepsilon|x|^2 \tag{4.41}$$

for any $(t,x)\in[0,s]\times\mathbb{R}^d$. We now try to derive the upper bound for $m^s_{\delta,\varepsilon}$ and $m_{\lambda,\delta,\varepsilon}$.

Since $[u(s,\cdot)]_1\leq M$ or $[\bar{u}(s,\cdot)]_1\leq M$, without loss of generality we assume $[\bar{u}(s,\cdot)]_1\leq M$, then $[\bar{v}(s,\cdot)]_1\leq Me^s$ and for any $x,y\in\mathbb{R}^d$

$$\begin{aligned}\psi^i(s,x,y) &\leq v_i(s,x) - \bar{v}_i(s,y) - \delta|x-y|^2\\ &\leq \sup_i|(v_i-\bar{v}_i)^+(s,\cdot)|_0 + Me^s|x-y| - \delta|x-y|^2\\ &\leq \sup_i|(v_i-\bar{v}_i)^+(s,\cdot)|_0 + M^2e^{2s}\delta^{-1}/4,\end{aligned}$$

where the last second inequality follows from that $\sup_{r\geq 0}(Cr-\delta r^2) = C^2/4\delta$ for any $C,\delta>0$. Thus, we get the upper bound for $m^s_{\delta,\varepsilon}$:

$$m^s_{\delta,\varepsilon} \leq \sup_i|(v_i-\bar{v}_i)^+(s,\cdot)|_0 + C_1\delta^{-1}, \tag{4.42}$$

where $C_1 = M^2e^{2s}/4$.

On the other hand, we assume that $m_{\lambda,\delta,\varepsilon}>0$ and derive its (positive) upper bound. Of course this upper bound still holds for $m_{\lambda,\delta,\varepsilon}\leq 0$. Follow this assumption, we have $t_0<s$, since otherwise, $m_{\lambda,\delta,\varepsilon} = \sup_{\mathbb{R}^d\times\mathbb{R}^d}\psi(s,x,y) - m^s_{\delta,\varepsilon}\leq 0$. Then we can apply the parabolic version of Crandall-Ishii's Lemma (see Lemma 2.2.12) to get that there are $a,b\in\mathbb{R}$ and $X,Y\in\mathcal{S}^d$ such that $(a, D_x\phi(t_0,x_0,y_0), X)\in\bar{\mathcal{P}}^{2,+}v_{i_0}(t_0,x_0)$ and $(b, -D_y\phi(t_0,x_0,y_0), Y)\in\bar{\mathcal{P}}^{2,-}\bar{v}_{i_0}(t_0,y_0)$ with $a-b=\phi_t(t_0,x_0,y_0)$ and the following inequality holds

$$\begin{pmatrix} X & 0\\ 0 & -Y\end{pmatrix} \leq 3e^{\lambda(s-t_0)}\delta\begin{pmatrix} I & -I\\ -I & I\end{pmatrix} + 3\varepsilon\begin{pmatrix} I & 0\\ 0 & I\end{pmatrix}. \tag{4.43}$$

Since $\bar{v}$ is a supersolution of (4.40) and $\bar{v}_{i_0}(t_0, y_0) < \mathcal{M}_{i_0}\bar{v}(t_0, y_0)$, we have

$$i) \quad \bar{v}_{i_0}(t_0, y_0) - \bar{f}'(t_0, y_0) \geq 0$$

or

$$ii) \quad -b + \bar{v}_{i_0}(t_0, y_0) + \sup_{\alpha \in \mathcal{A}} \inf_{\beta \in \mathcal{B}} \bar{\mathcal{L}}'^{\alpha,\beta}_{i_0}(t_0, y_0, -D_y\phi(t_0, x_0, y_0), Y) \geq 0.$$

Similar, since v is a subsolution, we have

$$iii) \quad v_{i_0}(t_0, x_0) - f'(t_0, x_0) \leq 0$$

and

$$iv) \quad -a + v_{i_0}(t_0, x_0) + \sup_{\alpha \in \mathcal{A}} \inf_{\beta \in \mathcal{B}} \mathcal{L}'^{\alpha,\beta}_{i_0}(t_0, x_0, D_x\phi(t_0, x_0, y_0), X) \leq 0.$$

We now discuss the two cases $i)$ and $ii)$ separately and derive an upper bound for $m_{\lambda,\delta,\varepsilon}$ that holds in either case.

If $i)$ holds, then combining with $iii)$ yields

$$\begin{aligned} m_{\lambda,\delta,\varepsilon} &\leq \psi^{i_0}(t_0, x_0, y_0) \leq v_{i_0}(t_0, x_0) - \bar{v}_{i_0}(t_0, y_0) \leq f'(t_0, x_0) - \bar{f}'(t_0, y_0) \\ &= e^{t_0}(f(t_0, x_0) - \bar{f}(t_0, y_0)) \leq e^{s}\left(|f - \bar{f}|_0 + [f]_{\mathcal{C}^{1,1/2}}|x_0 - y_0|\right) \end{aligned} \tag{4.44}$$

If $ii)$ holds, then combining with $iv)$ and using the fact that $\sup\inf(A) - \sup\inf(B) \leq \sup\sup(A - B)$, we have

$$\begin{aligned} v_{i_0}(t_0, x_0) - \bar{v}_{i_0}(t_0, y_0) &\leq a - b + \sup_{\alpha,\beta}\Big\{\bar{\mathcal{L}}'^{\alpha,\beta}_{i_0}(t_0, y_0, -D_y\phi(t_0, x_0, y_0), Y) \\ &- \ \mathcal{L}'^{\alpha,\beta}_{i_0}(t_0, x_0, D_x\phi(t_0, x_0, y_0), X)\Big\} \\ &\leq \ -\lambda e^{\lambda(s-t_0)}\delta|x_0 - y_0|^2 \\ &+ \ \sup_{\alpha,\beta}\Big\{\frac{1}{2}\mathrm{tr}[\sigma_{i_0}\sigma_{i_0}^T(t_0, x_0)X - \bar{\sigma}_{i_0}\bar{\sigma}_{i_0}^T(t_0, y_0)Y] \\ &- \ \bar{b}_{i_0}(t_0, y_0) \cdot (2e^{\lambda(s-t_0)}\delta(x_0 - y_0) - 2\varepsilon y_0) \\ &+ \ b_{i_0}(t_0, x_0) \cdot (2e^{\lambda(s-t_0)}\delta(x_0 - y_0) + 2\varepsilon x_0) \\ &- \ \bar{L}'_{i_0}(t_0, y_0) + L'_{i_0}(t_0, x_0)\Big\}. \end{aligned} \tag{4.45}$$

By the inequality (4.43) and the fact that $(s+t)^2 \leq 2(s^2 + t^2)$ for $s, t \in \mathbb{R}$, we obtain

$$\begin{aligned} &\mathrm{tr}[\sigma_{i_0}\sigma_{i_0}^T(t_0, x_0)X - \bar{\sigma}_{i_0}\bar{\sigma}_{i_0}^T(t_0, y_0)Y] \\ &\leq 6e^{\lambda(s-t_0)}\delta\left\{|\sigma_{i_0}(t_0, x_0) - \bar{\sigma}_{i_0}(t_0, x_0)|^2 + |\bar{\sigma}_{i_0}(t_0, x_0) - \bar{\sigma}_{i_0}(t_0, y_0)|^2\right\} \\ &+ 3\varepsilon\left\{|\sigma_{i_0}(t_0, x_0)|^2 + |\bar{\sigma}_{i_0}(t_0, y_0)|^2\right\} \\ &\leq 6e^{\lambda(s-t_0)}\delta\left\{|\sigma_{i_0} - \bar{\sigma}_{i_0}|_0^2 + [\bar{\sigma}_{i_0}]_{2,1}^2|x_0 - y_0|^2\right\} \\ &+ 3\varepsilon\left\{|\sigma_{i_0}|_0^2 + |\bar{\sigma}_{i_0}|_0^2\right\}. \end{aligned}$$

Furthermore, we have the following estimates

$$\begin{aligned}
&(b_{i_0}(t_0,x_0)-\bar{b}_{i_0}(t_0,y_0))\cdot(x_0-y_0)\\
\le\ &\frac{1}{2}\left(|b_{i_0}(t_0,x_0)-\bar{b}_{i_0}(t_0,x_0)|^2+|x_0-y_0|^2\right)\\
+\ &|\bar{b}_{i_0}(t_0,x_0)-\bar{b}_{i_0}(t_0,y_0)||x_0-y_0|\\
\le\ &\frac{1}{2}\left(|b_{i_0}-\bar{b}_{i_0}|_0^2+|x_0-y_0|^2\right)+[\bar{b}_{i_0}]_{2,1}|x_0-y_0|^2,
\end{aligned}$$

$$\begin{aligned}
b_{i_0}(t_0,x_0)\cdot x_0 &\le |b_{i_0}(t_0,x_0)||x_0| \le ([b_{i_0}]_{2,1}|x_0|+|b_{i_0}(t_0,0)|)|x_0|\\
&\le \bar{C}(1+|x_0|)^2 \le 2\bar{C}(1+|x_0|^2),
\end{aligned}$$

and similarly

$$\bar{b}_{i_0}(t_0,y_0)\cdot y_0 \le 2\bar{C}(1+|y_0|^2),$$

$$\bar{v}_{i_0}(t_0,y_0)-v_{i_0}(t_0,x_0)\le -\psi(t_0,x_0,y_0)\le -m_{\lambda,\delta,\varepsilon},$$

and

$$L'_{i_0}(t_0,x_0)-\bar{L}'_{i_0}(t_0,y_0)\le |L'_{i_0}-\bar{L}'_{i_0}|_0+[L'_{i_0}]|_{2,1}|x_0-y_0|.$$

Plugging in all these estimates into inequality (4.45) yields

$$\begin{aligned}
m_{\lambda,\delta,\varepsilon} &\le v_{i_0}(t_0,x_0)-\bar{v}_{i_0}(t_0,y_0)\\
&\le 3e^{\lambda(s-t_0)}\delta\sup_{i,\alpha,\beta}\left\{|\sigma-\bar{\sigma}|_0^2+|b-\bar{b}|_0^2\right\}+e^s\sup_{i,\alpha,\beta}|L-\bar{L}|_0\\
&+(C_2-\lambda)e^{\lambda(s-t_0)}\delta|x_0-y_0|^2+e^s[L_{i_0}]|_{2,1}|x_0-y_0|+C_3(1+|x_0|^2+|y_0|^2)\varepsilon.
\end{aligned}\tag{4.46}$$

where $C_2=3\bar{C}^2+2\bar{C}+1$, $C_3=3\bar{C}^2+4\bar{C}$ are some positive constants.

Combine (4.46) and (4.44), we get an upper bound for $m_{\lambda,\delta,\varepsilon}$ for either case. Plugging this upper bound as well as (4.42) into (4.41), we have

$$\begin{aligned}
&v_i(t,x)-\bar{v}_i(t,x)\\
\le\ &\sup_i|(v_i-\bar{v}_i)^+(s,\cdot)|_0+3e^{\lambda(s-t_0)}\delta\sup_{i,\alpha,\beta}\left\{|\sigma-\bar{\sigma}|_0^2+|b-\bar{b}|_0^2\right\}\\
+\ &e^s\{\sup_{i,\alpha,\beta}|L-\bar{L}|_0+|f-\bar{f}|_0\}+(C_2-\lambda)e^{\lambda(s-t_0)}\delta|x_0-y_0|^2+2e^s\bar{C}|x_0-y_0|\\
+\ &C_1\delta^{-1}+C_3(1+|x_0|^2+|y_0|^2)\varepsilon+2\varepsilon|x|^2
\end{aligned}$$

Note that this estimate holds for any $\lambda,\delta,\varepsilon>0$. We then try to select appropriate value for them (or take limit) to draw our conclusion. Firstly we may choose $\lambda=C_2+1$ and follow again that $\sup_{r\ge 0}(Cr-\delta r^2)=C^2/4\delta$ to get rid of the $|x_0-y_0|$ term. Then, by standard arguments, we know that for any fixed λ and δ, $\lim_{\varepsilon\to 0}\varepsilon(|x_0|^2+|y_0|^2)=0$. By letting $\varepsilon\to 0$, we further get

$$\begin{aligned}
&v_i(t,x)-\bar{v}_i(t,x)\\
\le\ &\sup_i|(v_i-\bar{v}_i)^+(s,\cdot)|_0+e^s\{\sup_{i,\alpha,\beta}|L-\bar{L}|_0+|f-\bar{f}|_0\}\\
+\ &3e^{(C_2+1)(s-t_0)}\delta\sup_{i,\alpha,\beta}\left\{|\sigma-\bar{\sigma}|_0^2+|b-\bar{b}|_0^2\right\}+(C_1+e^{2s}\bar{C}^2e^{-(C_2+1)(s-t_0)})\delta^{-1}.
\end{aligned}$$

Note that $\min_{r>0}(ar + br^{-1}) = 2(ab)^{1/2}$ for any $a, b > 0$, we can choose the δ minimising the right hand side to get

$$v_i(t,x) - \bar{v}_i(t,x) \le \sup_i |(v_i - \bar{v}_i)^+(s,\cdot)|_0 + e^s\{\sup_{i,\alpha,\beta} |L - \bar{L}|_0 + |f - \bar{f}|_0\} + C_4 \sup_{i,\alpha,\beta} \left\{|\sigma - \bar{\sigma}|_0 + |b - \bar{b}|_0\right\}.$$

where $C_4 = 2(3C_1 e^{(C_2+1)s} + 3e^{2s}\bar{C}^2)^{1/2}$ and we used that $(s^2+t^2)^{1/2} \le |s| + |t|$ for any $s, t \in \mathbb{R}$ in the last inequality. Finally, the conclusion follows by back-substituting v and $\bar{v}$ by u and $\bar{u}$. ∎

Finally, by using the above continuous dependence result, we show that the bounded viscosity solution u of (4.33)-(4.34) is the unique bounded solution, and moreover, it admits some regularity results.

Theorem 4.3.5 *Suppose that Assumption 4.3.1 is satisfied. Then, there exists a unique viscosity solution u of (4.33)-(4.34), with $|u|_1 \le C$ depending only on T and $\bar{C}$.*

Proof. In this proof, we denote by C some constant depending only on T and $\bar{C}$. Proposition 4.3.2 gives the existence, and the continuous dependence result (4.39) implies directly uniqueness and the x-regularity. To proof the t-regularity, we follow the idea in Appendix A. Fix any (t, s) such that $0 \le t < s \le T$ and define the functions $U_i^\varepsilon := u_i(s,\cdot) * \rho_\varepsilon$ in $\mathbb{R}^d$ for $i \in \mathcal{I}$ and some $\varepsilon > 0$ that shall be decided later, where ρ_ε is the same mollifiers defined in (A.3). Then similarly we have

$$|U_i^\varepsilon - u_i(s,\cdot)|_0 \le C\varepsilon \quad \text{and} \quad |D_x^j U_i^\varepsilon|_0 \le C\varepsilon^{1-|j|}.$$

Let u^ε be the unique bounded solution of (4.33) in Q_s with terminal condition U^ε. The continuous dependence results (4.39) then implies that for any $i \in \mathcal{I}$,

$$|u_i^\varepsilon - u_i| \le C \sup_i |U_i^\varepsilon - u_i(s,\cdot)|_0 \le C\varepsilon \quad \text{in } \bar{Q}_s.$$

Next, for each $i \in \mathcal{I}$, define two functions $w_{\varepsilon,i}^\pm(t,x) := U_i^\varepsilon(x) \pm C_\varepsilon(s-t)$ in $\bar{Q}_s$, for some $C_\varepsilon = C(\varepsilon^{-1}+1)$. It then can be easily checked that, for $i \in \mathcal{I}$, the functions $w_{\varepsilon,i}^-$ and $w_{\varepsilon,i}^+$ are bounded subsolution and bounded supersolution of

$$-\partial_t w + \sup_{\alpha\in\mathcal{A}} \inf_{\beta\in\mathcal{B}} \mathcal{L}_i^{\alpha,\beta}(t,x,D_x w, D_x^2 w) = 0, \quad \text{in } Q_s.$$

Thus, the function $\bar{v} = (\bar{v}_1, \bar{v}_2, ..., \bar{v}_m)$ such that $\bar{v}_i = w_{\varepsilon,i}^+$, is a bounded supersolution of (4.33) in Q_s. Applying (4.39) for u^ε and $\bar{v}$ yields that,

$$u_i^\varepsilon(t,x) - w_{\varepsilon,i}^+(t,x) \le C \sup_i |(U_i^\varepsilon - \bar{v}_i(s,\cdot))^+|_0 = 0$$

which implies that for $i \in \mathcal{I}$,

$$u_i^\varepsilon(t,x) - U_i^\varepsilon(x) \le C_\varepsilon(s-t).$$

Now we construct a bounded subsolution of (4.33) in Q_s based on w_ε^-. Note that for $i \in \mathcal{I}$ and any $j \neq i$,

$$U_i^\varepsilon = u_i(s,\cdot) * \rho_\varepsilon \leq \Big(\mathcal{M}_i^k u(s,\cdot)\Big) * \rho_\varepsilon \leq u_j(s,\cdot) * \rho_\varepsilon + k = U_j^\varepsilon + k,$$

we have $U_i^\varepsilon \leq \mathcal{M}_i^k U^\varepsilon$, and thus $w_{\varepsilon,i}^- - \mathcal{M}_i^k w_\varepsilon^- \leq 0$ in Q_s. Furthermore, similar to (B.3) in Appendix B, we have that in Q_s, $w_{\varepsilon,i}^- - f \leq C\varepsilon$. Thus, the function $\underline{v} = w_\varepsilon^- - C\varepsilon$ is a bounded subsolution of (4.33) in Q_s. Applying (4.39) for $\underline{v}$ and u^ε yields that for $i \in \mathcal{I}$,

$$w_{\varepsilon,i}^-(t,x) - C\varepsilon - u_i^\varepsilon(t,x) \leq 0,$$

which implies that for $i \in \mathcal{I}$,

$$U_i^\varepsilon(x) - u_i^\varepsilon(t,x) \leq C_\varepsilon(s-t) + C\varepsilon.$$

In turn, we obtain that for $i \in \mathcal{I}$,

$$\begin{aligned}
&|u_i(t,x) - u_i(s,x)| \\
&\leq |u_i(t,x) - u_i^\varepsilon(t,x)| + |u_i^\varepsilon(t,x) - U_i^\varepsilon(x)| + |U_i^\varepsilon(x) - u_i(s,x)| \\
&\leq 2C\varepsilon + C_\varepsilon(s-t) + C\varepsilon \\
&\leq C\left(\varepsilon + \frac{(s-t)}{\varepsilon} + (s-t)\right).
\end{aligned}$$

We choose $\varepsilon = \sqrt{s-t}$ to minimize the right hand side and obtain that for $i \in \mathcal{I}$,

$$|u_i(t,x) - u_i(s,x)| \leq C\sqrt{s-t}.$$

This completes the proof that $|u|_1 \leq C$. ■

Chapter 5

A monotone scheme for G-equations I: the theoretical results

In the last two chapters, we are concerned with semilinear equations/variational inequalities. In this chapter, we consider a special class of fully nonlinear PDEs called G-equations, which arise often in the characterization of G-distributed random variables in a sublinear expectation space. For such equations, we propose a monotone approximation scheme which is constructed recursively based on a piecewise constant approximation of the viscosity solution to the G-equation. We establish in this chapter the theoretical convergence of the scheme and determine the convergence rate, using the comparison principles for both the scheme and the equation together with the same mollification procedure specified before. We will then introduce some applications of this monotone approximation scheme in the next chapter.

5.1 Introduction

The theory of G-expectations (see [51–54]) is a natural generalization of classical probability theory in the presence of Knightian uncertainty. That is, random outcomes are evaluated, not using a single probability measure, but using the supremum over a range of possibly mutually singular probably measures. One of the fundamental results in the theory is the celebrated central limit theorem, dubbed as *robust central limit theorem* by Peng in [54]. It provides a theoretical foundation for the widely used G-distributed random variables in nonlinear probability and statistics. The theorem was first proved in [52] by applying the regularity theory of fully nonlinear PDEs (see [41] and [66]) to G-equations, the latter of which characterize G-distributed random variables. However, no convergence rate was derived in [52]. The corresponding convergence rate was subsequently obtained in [60] and [30] using Stein's method and more recently in [45] using stochastic control method under different model assumptions. However, an explicit formula for the constant appearing in the convergence rate is still lacking.

In this chapter, we build a monotone approximation scheme for the G-equation, and determine its convergence rate by obtaining an *explicit* error bound between the approximate solution and the viscosity solution of the G-equation. *This will, in turn, provide the convergence rate for Peng's robust central limit theorem with an explicit bound of Berry-Esseen type.* We are the first in the literature to obtain the explicit bound, although we mention that [45], where a similar technique to ours is applied, should also be able to keep track of all the constants and obtain a similar type of explicit bound. We will discuss this with more details in the next chapter. Different from [30], [45] and [60], our method is analytical and is developed under the framework of the monotone approximation schemes for viscosity solutions. Thus, it unveils an intrinsic connection between the convergence analysis of numerical schemes in PDEs and the central limit theorem in probability. It also introduces new tools from the numerical analysis for viscosity solutions to the study of G-expectations and especially its robust central limit theorem.

Let us first introduce Peng's G-equation. Let $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ be a sublinear expectation space, supporting two d-dimensional random vectors X and Y. Recall that $\hat{\mathbb{E}}$ is a sublinear expectation if it satisfies *monotonicity, constant preserving, sub-additivity* and *positive homogeneity* properties (see Section 6.1 for further details). With the random vectors (X, Y) and the sublinear expectation $\hat{\mathbb{E}}$, we introduce the nonlinear function $G : \mathbb{R}^d \times \mathbb{S}(d) \to \mathbb{R}$ as

$$G(p, A) := \hat{\mathbb{E}} \left[\langle p, Y \rangle + \frac{1}{2} \langle AX, X \rangle \right], \tag{5.1}$$

for $(p, A) \in \mathbb{R}^d \times \mathbb{S}(d)$, where $\mathbb{S}(d)$ is the collection of all d-dimensional symmetric matrixes.

For $T \geq 1$, let $Q_T := (0, T] \times \mathbb{R}^d$. We consider the fully-nonlinear parabolic PDE defined on the parabolic domain Q_T,

$$\partial_t u - G(D_x u, D_x^2 u) = 0, \tag{5.2}$$

with initial condition

$$u|_{t=0} = g. \tag{5.3}$$

In [51–54], the PDE (5.2) is referred to as the G-equation, which is used to characterize G-distribution. More specifically, let (ξ, ζ) be a pair of G-distributed (see Definition 6.1.5) d-dimensional random vectors characterized by (5.2) under another sublinear expectation $\widetilde{\mathbb{E}}$ (possibly different from $\hat{\mathbb{E}}$). That is, the G-distributed random vectors (ξ, ζ) satisfying

$$\widetilde{\mathbb{E}} \left[\langle p, \zeta \rangle + \frac{1}{2} \langle A\xi, \xi \rangle \right] = G(p, A) \tag{5.4}$$

for $(p, A) \in \mathbb{R}^d \times \mathbb{S}(d)$, and that for $a, b \in \mathbb{R}^d$ and $(\tilde{\xi}, \tilde{\zeta})$ as an independent copy of (ξ, ζ), the following equality holds in distribution sense:

$$\left(a\xi + b\tilde{\xi}, a^2\zeta + b^2\tilde{\zeta} \right) \stackrel{d}{=} \left(\sqrt{a^2 + b^2}\xi, (a^2 + b^2)\zeta \right).$$

Moreover, ζ is called maximal distributed in the sense that there exists a bounded, closed and convex subset $\Gamma \subset \mathbb{R}^d$ such that

$$\widetilde{\mathbb{E}}[\psi(\zeta)] = \max_{q \in \Gamma} \psi(q), \tag{5.5}$$

for any continuous function ψ satisfying linear growth condition. Note that the existence of (ξ, ζ) is guaranteed by Proposition 4.2 of [52]. Then, it has been proved in Proposition 4.8 of [52] that (5.2)-(5.3) admits a unique viscosity solution u which admits the representation

$$u(t,x) = \widetilde{\mathbb{E}}[g(x + \sqrt{t}\xi + t\zeta)], \tag{5.6}$$

provided that the initial data g satisfies some regularity condition. However, it is not clear how to explicitly solve (5.2)-(5.3) in order to characterize the G-distributed random vectors (ξ, ζ) except for some special cases, so a numerical scheme for (5.2)-(5.3) is needed.

We then propose a numerical scheme to approximate the viscosity solution u of (5.2)-(5.3) by merely using the random vectors (X, Y) under $\hat{\mathbb{E}}$ as input. Note that (X, Y) could follow arbitrary distributions (one example is shown in section 6.4). For $\Delta \in (0,1)$, we introduce $u^{\Delta} : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ recursively as

$$u^{\Delta}(t,x) = \hat{\mathbb{E}}[u^{\Delta}(t-\Delta, x + \sqrt{\Delta}X + \Delta Y)]\mathbf{1}_{\{t \geq \Delta\}} + g(x)\mathbf{1}_{\{t<\Delta\}}. \tag{5.7}$$

Remark 5.1.1 *The above recursive approximation implies that, for any $n \in \mathbb{N}$ such that $n\Delta \leq T$ and $t \in [n\Delta, ((n+1)\Delta) \wedge T)$, $u^{\Delta}(t,\cdot)$ is a constant in t and is given by*

$$u^{\Delta}(t,\cdot) \equiv u^{\Delta}(n\Delta, \cdot), \tag{5.8}$$

and at time $n\Delta$, there is a jump of the size

$$u^{\Delta}(n\Delta,\cdot) - u^{\Delta}((n-1)\Delta,\cdot) = \hat{\mathbb{E}}[u^{\Delta}((n-1)\Delta, \cdot + \sqrt{\Delta}X + \Delta Y)] - u^{\Delta}((n-1)\Delta,\cdot).$$

The main result of this chapter is proving the convergence of u^{Δ} to u and determining its convergence rate by obtaining the error bounds between the approximate solution and the viscosity solution of the G-equation. Note that the error bounds we obtained are explicit in the sense that the constant C has an explicit formula. This helps obtain the convergence rate of robust central limit theorem with an explicit bound of Berry-Esseen type.

We impose the following assumptions throughout the following two chapters.

Assumption 5.1.2 *(i) The initial data $g : \mathbb{R}^d \to \mathbb{R}$ is bounded from below, and β-Hölder continuous for some $\beta \in (0,1]$,*

$$|g(x) - g(y)| \leq C_g |x-y|^{\beta},$$

for $x, y \in \mathbb{R}^d$.

(ii) The random vectors X and Y satisfy the moment conditions: $M_X^3 < \infty$ and $M_Y^2 < \infty$, with $M_\xi^p := \hat{\mathbb{E}}[|\xi|^p]$. Moreover, X has no mean uncertainty, i.e. $\hat{\mathbb{E}}[X] = \hat{\mathbb{E}}[-X] = 0$.

Remark 5.1.3 *Assumptions (i) and (ii) are standard in the (robust) central limit theorem literature. The regularity of the initial condition g implies the regularity of the viscosity solution u (see Lemma 5.2.2). The bounded from below property of g guarantees the Fatou's property of $\hat{\mathbb{E}}$ (see (5.20) or Lemma 2.6 in [16]), which will in turn be used to establish an upper bound for the approximation error (see (5.40) in section 5.4.2).*

On the other hand, the moment conditions on X and Y are commonly used in the classical central limit theorem and imply that $M_X^p < \infty$ and $M_Y^q < \infty$ for $0 < p < 3$ and $0 < q < 2$. In our setting, they are used to derive the consistency error estimates in section 5.3.

Finally, we emphasize that there are no independence assumptions made between X and Y. If X and Y are mutually independent, then either (X, Y) must be maximally distributed or one of them is null (see [34]). The possible dependency between X and Y will be useful when applying the proposed approximation scheme to the Black-Scholes-Barenblatt equation in section 6.4.

Under the above assumptions, we prove the following results about the convergence of u^Δ to u and the corresponding convergence rate.

Theorem 5.1.4 *Suppose that Assumption 5.1.2 is satisfied. Then, the following assertions hold.*

(i) (Convergence) The approximate solution $u^\Delta \to u$ as $\Delta \to 0$, (locally) uniformly in $\bar{Q}_T$.

(ii) (Degenerate case) For $\Delta \in (0,1)$, there exists a constant C depending only on T, C_g, β, M_X^3 and M_Y^2 such that

$$|u - u^\Delta| \leq C\Delta^{\beta/6} \text{ in } \bar{Q}_T.$$

Furthermore, if the dimension $d = 1$ and $T = 1$, then the constant C has an explicit formula

$$C = 2124 C_g \left[1 + (M_X^3)^{\frac{\beta}{3}} + (M_Y^2)^{\frac{\beta}{2}}\right] \left[1 + (M_X^3)^{\frac{2}{3}} + M_X^3 + (M_Y^2)^{\frac{1}{2}} + M_Y^2\right]. \tag{5.9}$$

(iii) (Non-degenerate case) Furthermore, if the second moment of the random vector X is non-degenerate, i.e.

$$\underline{\sigma}^2 := -\hat{\mathbb{E}}[-|X|^2] > 0,$$

and the initial data $g \in \mathcal{C}_b^1(\mathbb{R}^d)$, i.e. g is bounded and Lipschitz continuous, then there exists a constant $\alpha \in (0,1)$ depending on d, $\underline{\sigma}^2$ and M_X^2 such that $u \in \mathcal{C}_b^{1+\frac{\alpha}{2}, 2+\alpha}(Q_T)$. Moreover, for $\Delta \in (0,1)$, there exists a constant C depending only on T, C_g, α, $M_X^{2+\alpha}$ and M_Y^2 such that

$$|u - u^\Delta| \leq C\Delta^{\max\{\frac{\alpha}{2}, \frac{1}{6}\}} \text{ in } \bar{Q}_T.$$

Assertion (i) is proved in section 5.3.2, assertion (ii) is proved in sections 5.4.1 and 5.4.2, and assertion (iii) is proved in section 5.4.3. We prove them under a similar framework of monotone approximation schemes for viscosity solutions that we use throughout this thesis. Finally, the non-degenerate situation (iii) is proved as a special case of the general (possibly degenerate) situation established in (ii), together with an application of the regularity theory of fully nonlinear PDEs.

Although the procedure we apply to obtain the approximation error bounds in this chapter is similar to that in the last two chapters, there are several differences that are worth to note. First, the sublinear expectation $\hat{\mathbb{E}}$ and the possible dependency between random vectors X and Y bring in additional difficulties when applying the monotone scheme. Specifically, the regularity properties of (approximate) solutions and the consistency error estimates are derived under the framework of G-expectations (see section 5.2 and Proposition 5.3.1), where the four axioms of the sublinear expectation $\hat{\mathbb{E}}$ (see Definition 6.1.1) are used in an essential way (without any independency assumption between X and Y). Second, since there are no variable coefficients in equation 5.2, the shaking coefficients technique is not needed. We apply the standard mollification procedure as in Section 2.4.1 to construct a sequence of smooth supersolutions (see (5.35)) due to the concavity of equation 5.2. This in turn leads to the derivation of the lower error bound. Moreover, by establishing almost the same regularity property for the approximate solution u^{Δ} as for the viscosity solution u in Lemma 5.2.3, we are able to interchange the roles of the G-equation and its approximation scheme, and thus obtain a symmetric upper bound and lower bound for the approximation error which is not the case in the previous two chapters. Last but not least, we work out explicit formulae for all the constants in our estimates. This enables us to derive an *explicit* error bound for the first time, which has a nontrivial application to Peng's robust central limit theorem.

Some supplementary notations to Section 2.1 to incorporate lower boundedness are needed in this chapter. For a function $f : \Omega \subset \mathbb{R}^d \to \mathbb{R}$, we define a semi-norm

$$[f]_0 := \sup_{x \in \Omega} f^-(x),$$

where $f^-(x) := \max(-f(x), 0)$ is the negative part of $f(x)$. We then denote

- $\mathcal{C}_{lb}(\Omega)$: the space of lower bounded continuous real-valued functions on Ω such that $[f]_0 < \infty$.
- $\mathcal{C}_{lb}^{\delta}(\Omega)$: the space of lower bounded continuous real-valued functions on Ω such that $[f]_0 + [f]_{\mathcal{C}^\delta} < \infty$.
- $\mathcal{C}_{lb}^{\infty}(\Omega)$: the space of lower bounded smooth real-valued functions on Ω with bounded derivatives of any order.

For functions of both time and space, the corresponding lower bounded function spaces are defined in a similar way.

5.2 Regularity estimates

We establish the space and time regularity properties of both u and u^{Δ}, which are crucial for proving the convergence of u^{Δ} to u and determining its convergence rate in Section 5.4.

Lemma 5.2.1 *Suppose that Assumption 5.1.2(ii) is satisfied. Then,*

(i) $\widetilde{\mathbb{E}}[|\xi|^2] = \hat{\mathbb{E}}[|X|^2] = M_X^2$*,*

(ii) $\widetilde{\mathbb{E}}[|\zeta|^p] \leq \hat{\mathbb{E}}[|Y|]^p = (M_Y^1)^p$ *for* $p > 0$*,*

where (ξ, ζ) *is a pair of G-distributed random vectors characterized by (5.1) with the associated sublinear expectation* $\widetilde{\mathbb{E}}$*.*

Proof. Assertion (i) is obvious by combining (5.1) and (5.4), and letting $p = 0$, $A = 2I_d$. By (5.5), we further have

$$\hat{\mathbb{E}}[\langle p, Y\rangle] = G(p, 0) = \widetilde{\mathbb{E}}[\langle p, \zeta\rangle] = \max_{q\in\Gamma}\langle p, q\rangle$$

for any $p \in \mathbb{R}^d$. Then, for any $q \in \Gamma$,

$$\begin{aligned}|q|^2 \leq\ & \max_{q'\in\Gamma}\langle q, q'\rangle = \hat{\mathbb{E}}[\langle q, Y\rangle] \leq \hat{\mathbb{E}}[\max_{q\in\Gamma}\langle q, Y\rangle] \\ &= \hat{\mathbb{E}}\left[\hat{\mathbb{E}}[\langle y, Y\rangle]|_{y=Y}\right] \leq \hat{\mathbb{E}}\left[\hat{\mathbb{E}}[|y||Y|]\,|_{y=Y}\right] = \hat{\mathbb{E}}\left[|Y|\hat{\mathbb{E}}[|Y|]\right] = \hat{\mathbb{E}}\left[|Y|\right]^2.\end{aligned}$$

Thus, we obtain from (5.5) again that $\widetilde{\mathbb{E}}[|\zeta|^p] = \max_{q\in\Gamma}|q|^p \leq \hat{\mathbb{E}}[|Y|]^p$. ■

Lemma 5.2.2 *Suppose that Assumption 5.1.2 is satisfied. Then, for any* $t, s \in [0, T]$ *and* $x, y \in \mathbb{R}^d$*,*

(i) $|u(t, x) - u(t, y)| \leq C_g|x - y|^\beta$*.*

(ii) $|u(s, x) - u(t, x)| \leq C_g K_0|s - t|^{\beta/2}$*, where the constant* K_0 *is defined as*

$$K_0 := e^{\frac{\beta T}{2}}[(M_X^2)^{\frac{\beta}{2}} + (M_Y^2)^{\frac{\beta}{2}}]. \tag{5.10}$$

Proof. Assertion (i) is a direct consequence of the representation formula (5.6), the sub-additivity of $\widetilde{\mathbb{E}}$ and the Hölder continuity of g.

To prove (ii), we may assume $t \leq s$. Note that the semigroup property of u implies that

$$u(s, x) = \widetilde{\mathbb{E}}[u(t, x + \sqrt{s-t}\xi + (s-t)\zeta)]. \tag{5.11}$$

In turn, the sub-additivity of $\widetilde{\mathbb{E}}$ and (i) yield

$$\begin{aligned}|u(s,x) - u(t,x)| &\leq \widetilde{\mathbb{E}}[|u(t, x + \sqrt{s-t}\xi + (s-t)\zeta) - u(t,x)|] \\ &\leq \widetilde{\mathbb{E}}[|u(t, x + \sqrt{s-t}\xi + (s-t)\zeta) - u(t, x + \sqrt{s-t}\xi)|] \\ &\quad + \widetilde{\mathbb{E}}[|u(t, x + \sqrt{s-t}\xi) - u(t,x)|] \\ &\leq \widetilde{\mathbb{E}}[C_g|(s-t)\zeta|^\beta] + \widetilde{\mathbb{E}}[C_g|\sqrt{s-t}\xi|^\beta] \\ &= C_g(\widetilde{\mathbb{E}}[|\xi|^\beta] + |s-t|^{\beta/2}\widetilde{\mathbb{E}}[|\zeta|^\beta])|s-t|^{\beta/2} \\ &\leq C_g((M_X^2)^{\beta/2} + |s-t|^{\beta/2}(M_Y^1)^\beta)|s-t|^{\beta/2}.\end{aligned}$$

where we used Lemma 5.2.1 and the fact that $\widetilde{\mathbb{E}}[|\xi|^\beta] \leq \widetilde{\mathbb{E}}[|\xi|^2]^{\beta/2}$ in the last inequality. The conclusion then follows from the inequalities

$$(M_X^2)^{\beta/2} + |s-t|^{\beta/2}(M_Y^1)^\beta \leq (M_X^2)^{\beta/2} + T^{\beta/2}(M_Y^2)^{\beta/2} \leq K_0.$$

■

Lemma 5.2.3 *Suppose that Assumption 5.1.2 is satisfied. Then, for any* $\Delta \in (0,1)$*,* $t, s \in [0,T]$ *and* $x, y \in \mathbb{R}^d$*,*

(i) $|u^{\Delta}(t,x) - u^{\Delta}(t,y)| \le C_g|x-y|^{\beta}$.

(ii) $|u^{\Delta}(s,x) - u^{\Delta}(t,x)| \le \sqrt{3}C_g K_0(|s-t|^{\beta/2} + \Delta^{\beta/2})$*, where the constant* K_0 *is given in (5.10).*

Proof. We first establish the estimate (i) using induction. It is clear that the estimate holds for $t \in [0,\Delta)$. In general, suppose the estimate holds for $t \in [(n-1)\Delta, n\Delta)$ with $n\Delta \le T$. Then, for $t \in [n\Delta, ((n+1)\Delta) \wedge T)$, the sub-additivity of $\hat{\mathbb{E}}$ yields

$$
\begin{aligned}
& |u^{\Delta}(t,x) - u^{\Delta}(t,y)| \\
= & \left|\hat{\mathbb{E}}[u^{\Delta}(t-\Delta, x+\sqrt{\Delta}X + \Delta Y)] - \hat{\mathbb{E}}[u^{\Delta}(t-\Delta, y+\sqrt{\Delta}X + \Delta Y)]\right| \\
\le & \hat{\mathbb{E}}\left|u^{\Delta}(t-\Delta, x+\sqrt{\Delta}X + \Delta Y) - u^{\Delta}(t-\Delta, y+\sqrt{\Delta}X + \Delta Y)\right| \\
\le & \hat{\mathbb{E}}[C_g|x-y|^{\beta}] = C_g|x-y|^{\beta}.
\end{aligned}
$$

where we also used the constant preserving property in the last inequality.

To establish the time regularity for u^{Δ} in (ii), we divide its proof into four steps.

Step 1. We lift the Hölder exponent β to 2 in the estimate (i). Note that the Young's inequality implies that

$$xy \le \frac{\beta}{2}x^{\frac{2}{\beta}} + \frac{2-\beta}{2}y^{\frac{2}{2-\beta}},\ x, y \ge 0.$$

In turn, for $\alpha \ge 0$ and $\varepsilon > 0$, let $x = \alpha^{\beta}$ and $y = \frac{1}{\varepsilon}$, and we have

$$\alpha^{\beta} \le \frac{\beta}{2}\varepsilon\alpha^2 + \frac{2-\beta}{2}\varepsilon^{\frac{-\beta}{2-\beta}},$$

Hence, it follows from (i) that

$$u^{\Delta}(t,x) \le u^{\Delta}(t,y) + a|x-y|^2 + b,\ x, y \in \mathbb{R}^d, \tag{5.12}$$

where $a := C_g\frac{\beta}{2}\varepsilon$ and $b := C_g\frac{2-\beta}{2}\varepsilon^{\frac{-\beta}{2-\beta}}$.

Step 2. Define $T_{\Delta} := \{k\Delta : k \in \mathbb{N}\}$. Then, for $\tau \in [0,T) \cap T_{\Delta}$ and $k \in \mathbb{N}$ such that $\tau + k\Delta \le T$, we aim to show that

$$u^{\Delta}(\tau + k\Delta, x) \le u^{\Delta}(\tau, y) + a(1+\Delta)^k|x-y|^2 + aNe^T k\Delta + b, \tag{5.13}$$

with a and b given in (5.12) and $N := 2M_X^2 + 3M_Y^2$. Indeed, it is clear that (5.13) holds for $k=0$. Suppose (5.13) holds for some $k \in \mathbb{N}$, then,

$$
\begin{aligned}
u^{\Delta}(\tau + (k+1)\Delta, x) &= \hat{\mathbb{E}}[u^{\Delta}(\tau + k\Delta, x + \sqrt{\Delta}X + \Delta Y)] \\
&\le u^{\Delta}(\tau, y) + a(1+\Delta)^k\hat{\mathbb{E}}[|x - y + \sqrt{\Delta}X + \Delta Y|^2] \\
&\quad + aNe^T k\Delta + b.
\end{aligned} \tag{5.14}
$$

For the sublinear expectation on the RHS of (5.14), we have

$$
\begin{aligned}
&\hat{\mathbb{E}}[|x-y+\sqrt{\Delta}X+\Delta Y|^2] \\
&\le |x-y|^2 + 2\hat{\mathbb{E}}[\langle x-y, \sqrt{\Delta}X\rangle] + 2\Delta\hat{\mathbb{E}}[\langle x-y, Y\rangle] + \Delta M_X^2 + \Delta^2 M_Y^2 + 2\Delta^{\frac{3}{2}}\hat{\mathbb{E}}[\langle X, Y\rangle].
\end{aligned}
$$

Since X has no mean uncertainty (cf. Assumption 5.1.2(ii)), it follows that $\hat{\mathbb{E}}[\langle x-y, \sqrt{\Delta}X\rangle] = 0$. Furthermore, since $2\langle x-y, Y\rangle \le |x-y|^2 + |Y|^2$ and $2\langle X, Y\rangle \le |X|^2 + |Y|^2$,

$$
\hat{\mathbb{E}}[|x-y+\sqrt{\Delta}X+\Delta Y|^2] \le (1+\Delta)|x-y|^2 + \Delta(2M_X^2 + 3M_Y^2) = (1+\Delta)|x-y|^2 + \Delta N. \tag{5.15}
$$

Combining (5.14)-(5.15) and the fact that $(1+\Delta)^k \le (1+\Delta)^{T/\Delta} \le e^T$, we have

$$
u^\Delta(\tau+(k+1)\Delta, x) \le u^\Delta(\tau, y) + a(1+\Delta)^{k+1}|x-y|^2 + aNe^T(k+1)\Delta + b,
$$

which shows that (5.13) also holds for $(k+1)$.

Step 3. We show that the estimate (ii) holds on $t, s \in [0,T) \cap T_\Delta$. Indeed, taking $y = x$ in (5.13), we obtain

$$
u^\Delta(\tau+k\Delta, x) \le u^\Delta(\tau, x) + C_g\frac{\beta}{2}\varepsilon Ne^T k\Delta + C_g\frac{2-\beta}{2}\varepsilon^{\frac{-\beta}{2-\beta}},
$$

for any $\varepsilon > 0$. Minimizing the RHS of the above inequality over ε then yields

$$
u^\Delta(\tau+k\Delta, x) \le u^\Delta(\tau, x) + C_g(Ne^T)^{\frac{\beta}{2}}(k\Delta)^{\frac{\beta}{2}}.
$$

Step 4. In general, for $s, t \in [0,T]$, let δ_s, $\delta_t \in [0,\Delta)$ such that $s-\delta_s$, $t-\delta_t \in T_\Delta$. Then, from (5.8) we have

$$
\begin{aligned}
u^\Delta(s,x) = u^\Delta(s-\delta_s, x) &\le u^\Delta(t-\delta_t, x) + C_g(Ne^T)^{\frac{\beta}{2}}(s-t-\delta_s+\delta_t)^{\beta/2} \\
&\le u^\Delta(t,x) + C_g(Ne^T)^{\frac{\beta}{2}}((s-t)^{\beta/2} + \Delta^{\beta/2}).
\end{aligned}
$$

Similarly, we also have

$$
u^\Delta(t,x) \le u^\Delta(s,x) + C_g(Ne^T)^{\frac{\beta}{2}}((s-t)^{\beta/2} + \Delta^{\beta/2}),
$$

from which we then conclude by observing that $(Ne^T)^{\frac{\beta}{2}} \le \sqrt{3}K_0$. ■

Remark 5.2.4 *Note that u^Δ is a piecewise constant approximation of u, so it is not continuous in time (with jumps at the partition points $\tau \in T_\Delta$). The discontinuity leads to the additional term $\Delta^{\beta/2}$ in the time regularity of u^Δ. Such type of time regularity property also appears in Lemma 2.2 of [45] in a stochastic control setting. Our regularity result could be regarded as a generalization of [45] to the sublinear expectation setting.*

5.3 A monotone approximation scheme for the G-equation

The proof of Theorem 5.1.4 is based on the monotone schemes for viscosity solutions. Similar to the previous two chapters, we first rewrite the recursive approximation (5.7) as a monotone scheme, and then derive its consistency error estimates.

Recall that $\mathcal{C}_{lb}(\mathbb{R}^d)$ is the space of lower bounded continuous functions on $\mathbb{R}^d$. We define a forward operator on $\mathcal{C}_{lb}(\mathbb{R}^d)$ as

$$\mathbf{S}(\Delta)\psi(x) = \hat{\mathbb{E}}[\psi(x + \sqrt{\Delta}X + \Delta Y)], \;\; \psi \in \mathcal{C}_{lb}(\mathbb{R}^d).$$

Then, from the four axioms of the sublinear expectation $\hat{\mathbb{E}}$ (see Definition 6.1), we immediately deduce that the forward operator $\mathbf{S}(\Delta)$ satisfies

(i) *(Monotonicity)* For any $\psi' \in \mathcal{C}_{lb}(\mathbb{R}^d)$ with $\psi' \geq \psi$,

$$\mathbf{S}(\Delta)\psi' \geq \mathbf{S}(\Delta)\psi. \tag{5.16}$$

(ii) *(Constant preserving)* For any $c \in \mathbb{R}$,

$$\mathbf{S}(\Delta)(\psi + c) = \mathbf{S}(\Delta)\psi + c. \tag{5.17}$$

(iii) *(Sub-additivity)* For any $\psi' \in \mathcal{C}_{lb}(\mathbb{R}^d)$,

$$\mathbf{S}(\Delta)(\psi' + \psi) \leq \mathbf{S}(\Delta)\psi' + \mathbf{S}(\Delta)\psi. \tag{5.18}$$

(iv) *(Positive homogeneity)* For any $\lambda \geq 0$,

$$\mathbf{S}(\Delta)(\lambda\psi) = \lambda\mathbf{S}(\Delta)\psi. \tag{5.19}$$

Note that (iii) and (iv) imply that $\mathbf{S}(\Delta)\psi$ is convex in ψ. On the other hand, the lower boundedness of ψ guarantee the Fatou's property (see Lemma 2.6 in [16]): Let $\psi_n \in \mathcal{C}_{lb}(\mathbb{R}^d)$ converges uniformly to ψ, then

$$\mathbf{S}(\Delta)\psi(x) \leq \liminf_{n} \mathbf{S}(\Delta)\psi_n(x). \tag{5.20}$$

The following error estimates play a vital rule to derive the consistency error estimates for the monotone approximation scheme introduced in section 5.3.1 (see Proposition 5.3.3(iii)).

Proposition 5.3.1 *Suppose that Assumption 5.1.2(ii) is satisfied. For $\Delta \in (0,1)$, define*

$$\mathcal{E}(\Delta, \psi) := \left| \frac{\mathbf{S}(\Delta)\psi - \psi}{\Delta} - G(D\psi, D^2\psi) \right|_0. \tag{5.21}$$

(i) If $\psi \in \mathcal{C}_b^{2+\alpha}(\mathbb{R}^d)$ for some $\alpha \in (0,1)$, then

$$\mathcal{E}(\Delta, \psi) \leq \;\Delta^{\alpha/2}[D^2\psi]_{\mathcal{C}^\alpha} M_X^{2+\alpha} + \sqrt{\Delta}|D^2\psi|_0(M_X^2 + M_Y^2).$$

(ii) If $\psi \in \mathcal{C}_{lb}^{\infty}(\mathbb{R}^d)$, then

$$\mathcal{E}(\Delta, \psi) \leq \;\sqrt{\Delta}|D^3\psi|_0 M_X^3 + \sqrt{\Delta}|D^2\psi|_0(M_X^2 + M_Y^2).$$

Proof. We only consider the case $d = 1$, since the general case follows along similar albeit more complicated arguments. Note that for any $x \in \mathbb{R}$,

$$\begin{aligned}
&\mathbf{S}(\Delta)\psi(x) - \psi(x) - \Delta G(D\psi(x), D^2\psi(x)) \\
&\leq \hat{\mathbb{E}}[\psi(x + \sqrt{\Delta}X + \Delta Y) - \psi(x) - \Delta D\psi(x)Y - \frac{1}{2}\Delta D^2\psi(x)X^2] \\
&\leq \hat{\mathbb{E}}[\psi(x + \sqrt{\Delta}X) - \psi(x) - \frac{1}{2}\Delta D^2\psi(x)X^2] \\
&\quad + \hat{\mathbb{E}}[\psi(x + \sqrt{\Delta}X + \Delta Y) - \psi(x + \sqrt{\Delta}X) - \Delta D\psi(x)Y] := (I) + (II).
\end{aligned}$$

Next, we obtain upper bounds for terms (I) and (II). To this end, Taylor's expansion and the assumption that $\hat{\mathbb{E}}[X] = \hat{\mathbb{E}}[-X] = 0$ yield

$$\begin{aligned}
(I) &= \hat{\mathbb{E}}\left[\sqrt{\Delta}D\psi(x)X + \int_x^{x+\sqrt{\Delta}X}\int_x^s (D^2\psi(u) - D^2\psi(x))duds\right] \\
&= \hat{\mathbb{E}}\left[\int_x^{x+\sqrt{\Delta}X}\int_x^s (D^2\psi(u) - D^2\psi(x))duds\right].
\end{aligned}$$

In case (i), $|D^2\psi(u) - D^2\psi(x)| \leq [D^2\psi]_{\mathcal{C}^\alpha}|u - x|^\alpha$, thus,

$$\begin{aligned}
(I) &\leq [D^2\psi]_{\mathcal{C}^\alpha}\hat{\mathbb{E}}\left[\int_x^{x+\sqrt{\Delta}X}\int_x^s |u - x|^\alpha duds\right] \\
&\leq [D^2\psi]_{\mathcal{C}^\alpha}\hat{\mathbb{E}}\left[\Delta^{1+\frac{\alpha}{2}}|X|^{2+\alpha}/(1+\alpha)(2+\alpha)\right] \leq \Delta^{1+\frac{\alpha}{2}}[D^2\psi]_{\mathcal{C}^\alpha}M_X^{2+\alpha}.
\end{aligned}$$

In case (ii), $|D^2\psi(u) - D^2\psi(x)| \leq |D^3\psi|_0|u - x|$, thus,

$$\begin{aligned}
(I) &= |D^3\psi|_0\hat{\mathbb{E}}\left[\int_x^{x+\sqrt{\Delta}X}\int_x^s |u - x|duds\right] \\
&\leq |D^3\psi|_0\hat{\mathbb{E}}\left[\Delta^{\frac{3}{2}}|X|^3/6\right] \leq \Delta^{\frac{3}{2}}|D^3\psi|_0 M_X^3.
\end{aligned}$$

Regarding term (II), for both cases (i) and (ii), we have

$$\begin{aligned}
(II) &= \hat{\mathbb{E}}\left[\int_{x+\sqrt{\Delta}X}^{x+\sqrt{\Delta}X+\Delta Y}(D\psi(s) - D\psi(x))ds\right] \\
&= \hat{\mathbb{E}}\left[\int_{x+\sqrt{\Delta}X}^{x+\sqrt{\Delta}X+\Delta Y}\int_x^s D^2\psi(u)duds\right] \\
&\leq |D^2\psi|_0\hat{\mathbb{E}}\left[\left|\frac{(\sqrt{\Delta}X + \Delta Y)^2 - (\sqrt{\Delta}X)^2}{2}\right|\mathbf{1}_{\{\sqrt{\Delta}X(\sqrt{\Delta}X+\Delta Y)\geq 0\}}\right. \\
&\quad \left. + \frac{(\sqrt{\Delta}X + \Delta Y)^2 + (\sqrt{\Delta}X)^2}{2}\mathbf{1}_{\{\sqrt{\Delta}X(\sqrt{\Delta}X+\Delta Y)< 0\}}\right] \\
&\leq |D^2\psi|_0\hat{\mathbb{E}}\left[(\Delta Y)^2/2 + \Delta^{\frac{3}{2}}|X||Y|\right] \\
&\leq |D^2\psi|_0\hat{\mathbb{E}}\left[(\Delta Y)^2/2 + \Delta^{\frac{3}{2}}(|X|^2 + |Y|^2)/2\right] \leq \Delta^{\frac{3}{2}}|D^2\psi|_0(M_X^2 + M_Y^2)
\end{aligned}$$

Combine the two estimates for terms (I) and (II), we obtain, for any $x \in \mathbb{R}$, that

$$\frac{\mathbf{S}(\Delta)\psi(x)-\psi(x)}{\Delta}-G(D\psi(x),D^2\psi(x))\le\Delta^{\alpha/2}[D^2\psi]_{\mathcal{C}^\alpha}M_X^{2+\alpha}+\sqrt{\Delta}|D^2\psi|_0(M_X^2+M_Y^2),$$

in case (i), and that

$$\frac{\mathbf{S}(\Delta)\psi(x)-\psi(x)}{\Delta}-G(D\psi(x),D^2\psi(x))\le\sqrt{\Delta}|D^3\psi|_0M_X^3+\sqrt{\Delta}|D^2\psi|_0(M_X^2+M_Y^2),$$

in case (ii). Similarly, we obtain lower bounds of $\mathcal{E}(\Delta,\psi)$, and this completes the proof. ■

5.3.1 The monotone approximation scheme

For $\Delta\in(0,1)$, we let $Q_T^\Delta:=(\Delta,T]\times\mathbb{R}^d$. Then, based on (5.7) and $\mathbf{S}(\Delta)$, we introduce the approximation scheme as

$$\begin{cases} S(\Delta,x,u^\Delta(t,x),u^\Delta(t-\Delta,\cdot))=0 & \text{in } \bar{Q}_T^\Delta,\\ u^\Delta(t,x)=g(x) & \text{in } \bar{Q}_T\backslash\bar{Q}_T^\Delta,\end{cases}\tag{5.22}$$

where $S:(0,1)\times\mathbb{R}^d\times\mathbb{R}\times\mathcal{C}_{lb}(\mathbb{R}^n)\to\mathbb{R}$ is defined by

$$S(\Delta,x,p,v)=\frac{p-\mathbf{S}(\Delta)v(x)}{\Delta}.\tag{5.23}$$

Remark 5.3.2 *The monotone approximation scheme (5.22) is a generalization of (2.40) to the sublinear expectation setting, with a 'first order term' variable Y added.*

From the properties of the forward operator $\mathbf{S}(\Delta)$ and Proposition 5.3.1, we obtain the following key properties of the approximation scheme (5.22).

Proposition 5.3.3 *Suppose that Assumption 5.1.2(ii) is satisfied. Then, the following properties hold for the approximation scheme $S(\Delta,x,p,v)$ given in (5.22).*

(i) (Monotonicity) For any $c_1,c_2\in\mathbb{R}$, and any function $u\in\mathcal{C}_{lb}(\mathbb{R}^n)$ with $u\le v$,

$$S(\Delta,x,p+c_1,u+c_2)\ge S(\Delta,x,p,v)+\frac{c_1-c_2}{\Delta}.$$

(ii) (Concavity) $S(\Delta,x,p,v)$ is concave in (p,v).

(iii) (Consistency) (a) If $\psi\in\mathcal{C}_b^{1+\frac{\alpha}{2},2+\alpha}(Q_T)$ for some $\alpha\in(0,1)$, then in Q_T^Δ,

$$\begin{aligned}&|\partial_t\psi-G(D_x\psi,D_x^2\psi)-S(\Delta,x,\psi,\psi(t-\Delta,\cdot))|\\ &\le K_\alpha\left(\Delta^{\alpha/2}\left([D_x^2\psi]_{\mathcal{C}^{\alpha/2,\alpha}}+[\partial_t\psi]_{\mathcal{C}^{\alpha/2,\alpha}}\right)+\sqrt{\Delta}|D_x^2\psi|_0+\Delta\left(|\partial_tD_x^2\psi|_0+|\partial_tD_x\psi|_0\right)\right),\end{aligned}\tag{5.24}$$

where the constant K_α is given by

$$K_\alpha:=1+M_Y^1+M_Y^2+M_X^2+M_X^{2+\alpha}.\tag{5.25}$$

(b) If $\psi \in \mathcal{C}_{lb}^{\infty}(Q_T)$, then in Q_T^{Δ},

$$
\begin{aligned}
&|\partial_t\psi - G(D_x\psi, D_x^2\psi) - S(\Delta, x, \psi, \psi(t-\Delta, \cdot))| \\
\le\ & K_1 \left(\sqrt{\Delta}\left(|D_x^3\psi|_0 + |D_x^2\psi|_0\right) + \Delta\left(|\partial_t^2\psi|_0 + |\partial_t D_x^2\psi|_0 + |\partial_t D_x\psi|_0\right)\right), \quad (5.26)
\end{aligned}
$$

where the constant K_1 is given by

$$K_1 := 1 + M_Y^1 + M_Y^2 + M_X^2 + M_X^3 \qquad (5.27)$$

Proof. Parts (i)-(ii) are immediate, so we only prove (iii). To this end, we split the consistency error into three parts. Specifically, for $(t, x) \in Q_T^{\Delta}$,

$$
\begin{aligned}
&|\partial_t\psi - G(D_x\psi, D_x^2\psi) - S(\Delta, x, \psi, \psi(t-\Delta, \cdot))| \\
\le\ & \mathcal{E}(\Delta, \psi(t-\Delta, \cdot)) + |\psi(t,x) - \psi(t-\Delta, x) - \Delta\partial_t\psi(t,x)|\Delta^{-1} \\
&+ |G(D_x\psi(t,x), D_x^2\psi(t,x)) - G(D_x\psi(t-\Delta, x), D_x^2\psi(t-\Delta, x))| \\
:=\ & (I) + (II) + (III),
\end{aligned}
$$

where $\mathcal{E}$ is defined in (5.21). Here we only consider the case (b); the case (a) only requires minor modification that is similar to the proof of Proposition 5.3.1(i). For term (I), Proposition 5.3.1 (ii) yields

$$\mathcal{E}(\Delta, \psi(t-\Delta, \cdot)) \le (M_X^3 + M_Y^2 + M_X^2)\sqrt{\Delta}\left(|D_x^3\psi|_0 + |D_x^2\psi|_0\right). \qquad (5.28)$$

For term (II), Taylor's expansion gives

$$
\begin{aligned}
&|\psi(t,x) - \psi(t-\Delta, x) - \Delta\partial_t\psi(t,x)|\Delta^{-1} \\
\le\ & \left|\int_{t-\Delta}^{t} (\partial_t\psi(s,x) - \partial_t\psi(t,x))\, ds\right|\Delta^{-1} \\
\le\ & \Delta^{-1}|\partial_t^2\psi|_0 \int_{t-\Delta}^{t} (t-s)\, ds \le \Delta|\partial_t^2\psi|_0. \qquad (5.29)
\end{aligned}
$$

Finally, for term (III), we have

$$
\begin{aligned}
&|G(D_x\psi(t,x), D_x^2\psi(t,x)) - G(D_x\psi(t-\Delta, x), D_x^2\psi(t-\Delta, x))| \\
\le\ & \hat{\mathbb{E}}[|D_x\psi(t,x) - D_x\psi(t-\Delta, x)||Y| + \frac{1}{2}|D_x^2\psi(t,x) - D_x^2\psi(t-\Delta, x)||X|^2] \\
\le\ & \Delta(M_Y^1|\partial_t D_x\psi|_0 + M_X^2|\partial_t D_x^2\psi|_0), \qquad (5.30)
\end{aligned}
$$

Combining estimates (5.28)-(5.30), we easily conclude. ■

Remark 5.3.4 *Due to the monotonicity property (i) in Proposition 5.3.3, the approximation scheme (5.22) is also referred to as the monotone (approximation) scheme in the sequel.*

The properties of the forward operator $\mathbf{S}(\Delta)$ also implies the following comparison principle for the monotone scheme (5.22). The proof is similar to that of Lemma 2.4.5, but under the sublinear expectation setting.

Proposition 5.3.5 *Suppose that Assumption 5.1.2(ii) is satisfied, and that u, $v \in \mathcal{C}_{lb}(\bar{Q}_T)$ are such that*

$$S(\Delta, x, u, u(t-\Delta, \cdot)) \le h_1 \text{ in } Q_T^\Delta,$$

$$S(\Delta, x, v, v(t-\Delta, \cdot)) \ge h_2 \text{ in } Q_T^\Delta,$$

for some h_1, $h_2 \in \mathcal{C}_{lb}(Q_T^\Delta)$. Then,

$$u - v \le \sup_{\bar{Q}_T \setminus Q_T^\Delta} (u-v)^+ + t \sup_{Q_T^\Delta} (h_1 - h_2)^+ \text{ in } \bar{Q}_T. \tag{5.31}$$

Proof. (5.31) holds obviously in $\bar{Q}_T \setminus Q_T^\Delta$. In Q_T^Δ, we have

$$u(t,x) \le \mathbf{S}(\Delta) u(t-\Delta, \cdot)(x) + \Delta h_1$$

and

$$v(t,x) \ge \mathbf{S}(\Delta) v(t-\Delta, \cdot)(x) + \Delta h_2$$

Combining the above two inequalities and using sub-additivity (5.18) then yield

$$(u-v)(t,x) \le \sup_{y \in \mathbb{R}} (u-v)(t-\Delta, y) + \Delta \sup_{Q_T^\Delta} (h_1 - h_2)^+.$$

The conclusion then follows by induction. ■

5.3.2 Convergence of the monotone approximation scheme

We prove Theorem 5.1.4(i) by showing the convergence of the approximate solution u^Δ to the viscosity solution u. It is based on the monotone schemes for viscosity solutions introduced by Barles-Souganidis in [5], where they show that any *monotone, stable* and *consistent* numerical scheme converges, provided that there exists a strong comparison principle for the equation (see Definition 2.3.1 and Theorem 2.3.2). However, our proof uses only a standard comparison principle (provided by Corollary 6.4 of [52]), thanks to the regularity results of u^Δ we obtained in Lemma 5.2.3.

Specifically, define the semi-relaxed limits of u^Δ by

$$\overline{u}(t,x) = \limsup_{\substack{(t',x') \to (t,x), \\ \Delta \to 0}} u^\Delta(t',x'); \quad \underline{u}(t,x) = \liminf_{\substack{(t',x') \to (t,x), \\ \Delta \to 0}} u^\Delta(t',x').$$

We show that $\overline{u}$ is a viscosity subsolution of (5.2)-(5.3). A symmetric argument will imply that $\underline{u}$ is a viscosity supersolution of (5.2), which proves that $\overline{u} = \underline{u} = u$, so u^Δ converges to u locally uniformly.

Let $\phi \in \mathcal{C}^\infty(\bar{Q}_T)$ and $(t_0, x_0) \in Q_T$ be such that

$$0 = (\overline{u} - \phi)(t_0, x_0) = \max_{(t',x')} (\overline{u} - \phi)(t', x').$$

By the definition of $\overline{u}$, there exists a sequence $\{(t_n, x_n, \Delta_n)\}_{n \ge 1}$ such that

$$(t_n, x_n, \Delta_n) \to (t_0, x_0, 0), \quad \text{and} \quad u^{\Delta_n}(t_n, x_n) \to \overline{u}(t_0, x_0).$$

Moreover, by extracting a subsequence if necessary, (t_n, x_n) is also the maximum point of $u^{\Delta_n} - \phi$:

$$\delta^{\Delta_n} := (u^{\Delta_n} - \phi)(t_n, x_n) = \max_{(t',x')}(u^{\Delta_n} - \phi)(t', x') \to 0.$$

Since $t_0 > 0$ and $\Delta \to 0$, we have $t_n > \Delta_n$ for large enough n. The monotonicity property (i) in Proposition 5.3.3 further implies that

$$\begin{aligned}
0 &= S(\Delta_n, x_n, u^{\Delta_n}(t_n, x_n), u^{\Delta_n}(t_n - \Delta_n, \cdot)) \\
&\geq S(\Delta_n, x_n, \phi(t_n, x_n) + \delta^{\Delta_n}, \phi(t_n - \Delta_n, \cdot) + \delta^{\Delta_n}) \\
&= \frac{\phi(t_n, x_n) - \mathbf{S}(\Delta_n)\phi(t_n - \Delta_n, \cdot)(x_n)}{\Delta_n}.
\end{aligned}$$

In turn, using the consistency property (iii) in Proposition 5.3.3 and letting $n \to \infty$, we obtain

$$\partial_t \phi(t_0, x_0) - G(D_x\phi(t_0, x_0), D_x^2\phi(t_0, x_0)) \leq 0.$$

Next, we show that $\overline{u}(0, x) = g(x)$ for $x \in \mathbb{R}^d$. Let $\{(t_n, x_n, \Delta_n)\}_{n\geq 1}$ be a sequence such that

$$(t_n, x_n, \Delta_n) \to (0, x, 0), \quad \text{and} \quad u^{\Delta_n}(t_n, x_n) \to \overline{u}(0, x).$$

Since $u^{\Delta_n}(s, x_n) = g(x_n)$ for $s \in \bar{Q}_T \backslash \bar{Q}_T^{\Delta_n}$, by the time regularity of u^Δ in Lemma 5.2.3(ii), we have

$$|u^{\Delta_n}(t_n, x_n) - u^{\Delta_n}(s, x_n)| \leq \sqrt{3}C_g K_0(|t_n - s|^{\beta/2} + \Delta_n^{\beta/2}).$$

Letting $s = 0$ and sending $n \to \infty$ yield that $|\bar{u}(0, x) - g(x)| = 0$, from which we conclude that $\overline{u}(\cdot, \cdot)$ is a viscosity subsolution of (5.2)-(5.3).

5.4 Convergence rate of the monotone approximation scheme

In this section, we prove Theorem 5.1.4(ii) by establishing the (uniform) convergence rate of the approximate solution u^Δ to the viscosity solution u, and keeping track of all the involved constants. We start with the approximation error in the first time interval $\bar{Q}_T \backslash Q_T^\Delta$, where $u^\Delta = g = u|_{t=0}$ except at $t = \Delta$. Therefore, the bound for the approximation error in this interval can be easily obtained by the regularity property of u in Lemmas 5.2.2. This is demonstrated in the following lemma.

Lemma 5.4.1 *Suppose that Assumption 5.1.2 is satisfied. Then, for $\Delta \in (0, 1)$,*

$$|u - u^\Delta| \leq 2C_g K_0 \Delta^{\beta/2} \ \text{in} \ \bar{Q}_T \backslash Q_T^\Delta, \tag{5.32}$$

where the constant K_0 is given in (5.10).

Proof. Since $u^\Delta = g = u|_{t=0}$ in $\bar{Q}_T \backslash \bar{Q}_T^\Delta$, we have, for $(t, x) \in \bar{Q}_T \backslash Q_T^\Delta$,

$$|u(t, x) - u^\Delta(t, x)| \leq |u(t, x) - u(0, x)| + |u^\Delta(\Delta, x) - u^\Delta(0, x)|\mathbf{1}_{\{t=\Delta\}}.$$

When $t = \Delta$, we further obtain

$$\begin{aligned}|u^\Delta(\Delta, x) - u^\Delta(0, x)| &\le \hat{\mathbb{E}}[|u^\Delta(0, x + \sqrt{\Delta}X + \Delta Y) - u^\Delta(0, x)|]\\ &= \hat{\mathbb{E}}[|g(x + \sqrt{\Delta}X + \Delta Y) - g(x)|]\\ &\le \hat{\mathbb{E}}[C_g|\sqrt{\Delta}X + \Delta Y|^\beta]\\ &\le C_g(M_X^\beta + \Delta^{\beta/2}M_Y^\beta)\Delta^{\beta/2}\\ &\le C_g((M_X^2)^{\beta/2} + (M_Y^2)^{\beta/2})\Delta^{\beta/2} \le C_g K_0 \Delta^{\beta/2}.\end{aligned}$$

The conclusion then follows from Lemma 5.2.2(ii). ■

5.4.1 Lower bound for the approximation error

For $u \in \mathcal{C}_{lb}^{\frac{\beta}{2},\beta}(\bar{Q}_T)$, we aim to derive a lower bound for the approximation error $u - u^\Delta$ within the whole domain $\bar{Q}_T$. To this end, for $\varepsilon \in (0, 1)$, we extend the domain of the G-equation (5.2) from Q_T to $Q_{T+\varepsilon^2} := (0, T + \varepsilon^2] \times \mathbb{R}^d$ and still denote the solution as u. Next, we regularize u by a standard mollification procedure: for $(t, x) \in \bar{Q}_T$, we define

$$u_\varepsilon(t, x) = u * \rho_\varepsilon(t, x) = \int_{-\varepsilon^2<\tau<0} \int_{|e|<\varepsilon} u(t - \tau, x - e)\rho_\varepsilon(\tau, e)ded\tau,$$

where the sequence of mollifiers ρ_ε are defined in (3.30). Lemma 5.2.2 implies that

$$|u(t, x) - u(s, y)| \le C_g \left[|x - y|^\beta + K_0|s - t|^{\beta/2}\right].$$

In turn, standard properties of mollifiers (see (2.28) and (2.29) in Section 2.4.1) imply that $u_\varepsilon \in \mathcal{C}_{lb}^\infty(\bar{Q}_T)$,

$$|u - u_\varepsilon|_0 \le C_g(1 + K_0)\varepsilon^\beta, \tag{5.33}$$

and, moreover, for positive integer i and multi-index j,

$$|\partial_t^i D_x^j u_\varepsilon|_0 \le C_g(1 + K_0)\varepsilon^{\beta-2i-|j|}||\partial_t^i D_x^j \rho||_1, \tag{5.34}$$

where the constant K_0 is given in (5.10) and

$$||\partial_t^i D_x^j \rho||_1 = \int_{B(0,1)} \left|\partial_t^i D_x^j \rho(\tau, e)\right| d(\tau, e) < \infty.$$

Lemma 5.4.2 *$u_\varepsilon(t, x)$ is a (classical) supersolution of (5.2) in Q_T, namely,*

$$\partial_t u_\varepsilon - G(D_x u_\varepsilon, D_x^2 u_\varepsilon) \ge 0. \tag{5.35}$$

Proof. We observe that the function $u(t - \tau, x - e)$ is still a viscosity solution of the G-equation (5.2) in Q_T for any $(\tau, e) \in (-\varepsilon^2, 0) \times B(0, \varepsilon)$. On the other hand, a Riemann sum approximation shows that there exists a sequence $\{I_n\}_{n\ge1} \in \mathcal{C}_{lb}(\bar{Q}_T)$ such that each I_n is a convex combination of the functions $u(\cdot - \tau, \cdot - e)$ for different $(\tau, e) \in (-\varepsilon^2, 0) \times B(0, \varepsilon)$ and that I_n converges uniformly to u_ε. Since the nonlinear term $G(p, X)$ is convex in p and X, each I_n becomes a supersolution of (5.2) in Q_T. Using the stability of viscosity solutions (Proposition 2.2.6), we deduce that $u_\varepsilon(t,x)$ is still a supersolution of (5.2) in Q_T. ■

We are now in a position to establish a lower bound for the approximation error.

Theorem 5.4.3 *Suppose that Assumption 5.1.2 is satisfied. Then, for $\Delta \in (0,1)$, there exists a constant C_{LB} depending only on T, C_g, β, M_X^3 and M_Y^2 such that*

$$u - u^\Delta \geq -C_{LB}\Delta^{\beta/6} \text{ in } \bar{Q}_T.$$

Moreover, the constant C_{LB} has an explicit formula $C_{LB} := C_g(1+K_0)\left(4 + K_1 C_\rho T\right)$ with the constants K_0, K_1 and C_ρ given in (5.10), (5.27) and (5.37), respectively.

Proof. Since $u_\varepsilon \in \mathcal{C}_{lb}^\infty(\bar{Q}_T)$ is smooth with bounded derivatives of any order, we substitute u_ε into the consistency error estimate (5.26) and use (5.35) and (5.34) to obtain

$$\begin{aligned}
& S(\Delta, x, u_\varepsilon(t,x), u_\varepsilon(t-\Delta,\cdot)) \\
\geq\; & -C_g(1+K_0)K_1 \\
& \times \Big[\sqrt{\Delta}\left(\varepsilon^{\beta-3}||D_x^3\rho||_1 + \varepsilon^{\beta-2}||D_x^2\rho||_1\right) \\
& +\Delta\left(\varepsilon^{\beta-4}(||\partial_t^2\rho||_1 + ||\partial_t D_x^2\rho||_1) + \varepsilon^{\beta-3}||\partial_t D_x\rho||_1\right)\Big] \\
\geq\; & -C_g(1+K_0)K_1 \\
& \times \Big[\sqrt{\Delta}\varepsilon^{\beta-3}\left(||D_x^3\rho||_1 + ||D_x^2\rho||_1\right) + \Delta\varepsilon^{\beta-4}\left(||\partial_t^2\rho||_1 + ||\partial_t D_x^2\rho||_1 + ||\partial_t D_x\rho||_1\right)\Big] \\
=:\; & -C_g(1+K_0)K_1 c(\beta,\varepsilon),
\end{aligned} \tag{5.36}$$

for $(t,x) \in Q_T^\Delta$, where the constants K_0 and K_1 are given in (5.10) and (5.26), respectively. The comparison principle in Proposition 5.3.5 then implies that in $\bar{Q}_T$,

$$u^\Delta - u_\varepsilon \leq \sup_{\bar{Q}_T \setminus Q_T^\Delta}(u^\Delta - u_\varepsilon)^+ + c(\beta,\varepsilon)TC_g(1+K_0)K_1.$$

Next, using (5.33), we further obtain

$$\begin{aligned}
u^\Delta - u &= (u_\varepsilon - u) + (u^\Delta - u_\varepsilon) \\
&\leq C_g(1+K_0)\varepsilon^\beta + \sup_{\bar{Q}_T \setminus Q_T^\Delta}(u^\Delta - u_\varepsilon)^+ + c(\beta,\varepsilon)TC_g(1+K_0)K_1 \\
&\leq \sup_{\bar{Q}_T \setminus Q_T^\Delta}(u - u^\Delta)^+ + 2C_g(1+K_0)\varepsilon^\beta + c(\beta,\varepsilon)TC_g(1+K_0)K_1 \text{ in } \bar{Q}_T.
\end{aligned}$$

By choosing $\varepsilon = \Delta^{1/6}$, we conclude that

$$\begin{aligned}
u^\Delta - u &\leq \sup_{\bar{Q}_T \setminus Q_T^\Delta}(u - u^\Delta)^+ + 2C_g(1+K_0)\Delta^{\beta/6} + c(\beta,\Delta^{1/6})TC_g(1+K_0)K_1 \\
&\leq C_g(1+K_0)\left(4 + K_1 C_\rho T\right)\Delta^{\beta/6} \quad \text{in } \bar{Q}_T,
\end{aligned}$$

where the last inequality follows from the estimate (5.32) in Lemma 5.4.1 and the fact that $c(\beta, \Delta^{1/6}) \le C_\rho \Delta^{\beta/6}$ with

$$C_\rho := ||D_x^3\rho||_1 + ||D_x^2\rho||_1 + ||\partial_t^2\rho||_1 + ||\partial_t D_x^2\rho||_1 + ||\partial_t D_x\rho||_1 < \infty. \quad (5.37)$$

■

Remark 5.4.4 *A typical example of ρ is given by*

$$\rho(t,x) = K\exp(-\frac{1}{1-|x|^2})\exp(-\frac{1}{1-(2t+1)^2})\mathbf{1}_{\{|x|<1,-1<t<0\}},$$

where K is given such that the mass of ρ is 1. One can always compute C_ρ using this example and in the one-dimension case, $C_\rho < 10^3 e^{-1}$. In turn, when $T=1$, it follows from the formulae for K_0 and K_1 (c.f. (5.10) and (5.27)) that

$$\begin{aligned} C_{LB} &\le C_g\left[1 + e^{\frac{\beta}{2}}((M_X^2)^{\frac{\beta}{2}} + (M_Y^2)^{\frac{\beta}{2}})\right]\left[4 + (1 + M_Y^1 + M_Y^2 + M_X^2 + M_X^3)\frac{10^3}{e}\right] \\ &\le 613C_g\left[1 + (M_X^3)^{\frac{\beta}{3}} + (M_Y^2)^{\frac{\beta}{2}}\right]\left[1 + (M_X^3)^{\frac{2}{3}} + M_X^3 + (M_Y^2)^{\frac{1}{2}} + M_Y^2\right]. \end{aligned}$$

5.4.2 Upper bound for the approximation error

To obtain an upper bound for the approximation error, we are not able to construct approximate smooth subsolutions of (5.2) due to the convexity of the function G. Instead, we interchange the roles of the G-equation (5.2) and the monotone scheme (5.22) (as in [33] and [39]). Note that we can do this thanks to the regularity results of u^Δ in Lemma 5.2.3.

To this end, for $\varepsilon \in (0,1)$, we extend the domain of the monotone scheme (5.22) from $\bar{Q}_T$ to $\bar{Q}_{T+\varepsilon^2} := [0, T+\varepsilon^2] \times \mathbb{R}^d$ and still denote the scheme solution as u^Δ. Then, using the same mollifiers ρ_ε as in section 5.4.1, we define, for $(t,x) \in \bar{Q}_T$,

$$u_\varepsilon^\Delta(t,x) = u^\Delta * \rho_\varepsilon(t,x) = \int_{-\varepsilon^2<\tau<0}\int_{|e|<\varepsilon} u^\Delta(t-\tau, x-e)\rho_\varepsilon(\tau,e)ded\tau.$$

The regularity property of u^Δ in Lemma 5.2.3 implies that

$$|u^\Delta(t,x) - u^\Delta(s,y)| \le C_g\left[|x-y|^\beta + \sqrt{3}K_0(|s-t|^{\beta/2} + \Delta^{\beta/2})\right].$$

Again, standard properties of mollifiers imply that $u_\varepsilon^\Delta \in \mathcal{C}_{lb}^\infty(\bar{Q}_T)$,

$$|u^\Delta - u_\varepsilon^\Delta|_0 \le C_g(1+\sqrt{3}K_0)(\varepsilon^\beta + \Delta^{\beta/2}), \quad (5.38)$$

and, moreover, for positive integer i and multi-index j,

$$|\partial_t^i D_x^j u_\varepsilon^\Delta|_0 \le C_g(1+\sqrt{3}K_0)\varepsilon^{-2i-|j|}(\varepsilon^\beta + \Delta^{\beta/2})||\partial_t^i D_x^j\rho||_1. \quad (5.39)$$

Next, let $\{I_n^\Delta\}_{n\ge1} \in \mathcal{C}_{lb}(\bar{Q}_T)$ be a sequence such that each I_n^Δ is a convex combination of the functions $u^\Delta(\cdot-\tau, \cdot-e)$ for different $(\tau, e) \in (-\varepsilon^2, 0) \times B(0,\varepsilon)$ and that I_n^Δ converges uniformly to u_ε^Δ. Since

$$S(\Delta, x, u^\Delta(t-\tau, x-e), u^\Delta(t-\tau-\Delta, \cdot-e)) = 0 \text{ in } \bar{Q}_T^\Delta,$$

for any $(\tau, e) \in (-\varepsilon^2, 0) \times B(0, \varepsilon)$, the concavity of the monotone scheme (cf. Proposition 5.3.3 (ii)) yields that for any $n \in \mathbb{N}$ and $(t, x) \in \bar{Q}_T^\Delta$,

$$S(\Delta, x, I_n^\Delta(t, x), I_n^\Delta(t - \Delta, \cdot)) \geq 0.$$

Since I_n^Δ is lower bounded, we use Fatou's property of the sublinear expectation $\hat{\mathbb{E}}$ (see (5.20)) to deduce that, for $(t, x) \in \bar{Q}_T^\Delta$,

$$\begin{aligned}
& S(\Delta, x, u_\varepsilon^\Delta(t, x), u_\varepsilon^\Delta(t - \Delta, \cdot)) \\
= \; & \left(u_\varepsilon^\Delta(t, x) - \hat{\mathbb{E}}[u_\varepsilon^\Delta(t - \Delta, x + \sqrt{\Delta} X + \Delta Y)] \right) \Delta^{-1} \\
\geq \; & \left(u_\varepsilon^\Delta(t, x) - \lim_{n \to \infty} \hat{\mathbb{E}}[I_n^\Delta(t - \Delta, x + \sqrt{\Delta} X + \Delta Y)] \right) \Delta^{-1} \\
= \; & \lim_{n \to \infty} S(\Delta, x, I_n^\Delta(t, x), I_n^\Delta(t - \Delta, \cdot)) \geq 0. \qquad (5.40)
\end{aligned}$$

We are now in a position to establish an upper bound for the approximation error.

Theorem 5.4.5 *Suppose that Assumption 5.1.2 is satisfied. Then, for $\Delta \in (0, 1)$, there exists a constant C_{UB} depending only on T, C_g, β, M_X^3 and M_Y^2 such that*

$$u - u^\Delta \leq C_{UB} \Delta^{\beta/6} \text{ in } \bar{Q}_T.$$

Moreover, the constant C_{UB} has an explicit formula $C_{UB} := 2\sqrt{3} C_{LB} = 2\sqrt{3} C_g (1 + K_0)(4 + K_1 C_\rho T)$ with the constants K_0, K_1 and C_ρ given in (5.10), (5.26) and (5.37), respectively.

Proof. We first consider the above error estimate in Q_T^Δ. Since $u_\varepsilon^\Delta \in \mathcal{C}_{lb}^\infty(\bar{Q}_T)$ is smooth with bounded derivatives of any order, we substitute u_ε^Δ into the consistency error estimate (5.26) and use (5.40) and (5.39) to obtain

$$\partial_t u_\varepsilon^\Delta - G(D_x u_\varepsilon^\Delta, D_x^2 u_\varepsilon^\Delta) \geq -C_g (1 + \sqrt{3} K_0) K_1 (\varepsilon^\beta + \Delta^{\beta/2}) c(0, \varepsilon)$$

for $(t, x) \in Q_T^\Delta$, where $c(0, \varepsilon)$ is defined in (5.36). Then, the function

$$\bar{v}(t, x) := u_\varepsilon^\Delta(t, x) + C_g (1 + \sqrt{3} K_0) K_1 (\varepsilon^\beta + \Delta^{\beta/2}) c(0, \varepsilon)(t - \Delta)$$

becomes a (classical) supersolution of the G-equation (5.2) in Q_T^Δ with initial condition $\bar{v}(\Delta, x) = u_\varepsilon^\Delta(\Delta, x)$. On the other hand, from (5.32) and (5.38), we know that

$$\underline{v}(t, x) := u(t, x) - C_g (1 + \sqrt{3} K_0)(\varepsilon^\beta + \Delta^{\beta/2}) - 2 C_g K_0 \Delta^{\beta/2}$$

is a (viscosity) solution of the G-equation (5.2), such that

$$\begin{aligned}
\underline{v}(\Delta, x) &= u(\Delta, x) - C_g (1 + \sqrt{3} K_0)(\varepsilon^\beta + \Delta^{\beta/2}) - 2 C_g K_0 \Delta^{\beta/2} \\
&= u(\Delta, x) - u^\Delta(\Delta, x) + u^\Delta(\Delta, x) - u_\varepsilon^\Delta(\Delta, x) + u_\varepsilon^\Delta(\Delta, x) \\
&\quad - 2 C_g K_0 \Delta^{\beta/2} - C_g (1 + \sqrt{3} K_0)(\varepsilon^\beta + \Delta^{\beta/2}) \leq u_\varepsilon^\Delta(\Delta, x) = \bar{v}(\Delta, x).
\end{aligned}$$

Thus, the comparison principle for the G-equation (see Corollary 6.4 in [52]) implies that $\underline{v} \le \bar{v}$ in $\bar{Q}_T^\Delta$, i.e.

$$\begin{aligned} u - u_\varepsilon^\Delta \le\ & C_g(1+\sqrt{3}K_0)(\varepsilon^\beta + \Delta^{\beta/2}) + 2C_g K_0 \Delta^{\beta/2} \\ & + C_g(1+\sqrt{3}K_0)K_1(\varepsilon^\beta + \Delta^{\beta/2})c(0,\varepsilon)(t-\Delta) \text{ in } \bar{Q}_T^\Delta. \end{aligned}$$

Finally, using the estimates (5.38) again, we obtain by choosing $\varepsilon = \Delta^{1/6}$ that

$$\begin{aligned} u - u^\Delta &= (u - u_\varepsilon^\Delta) + (u_\varepsilon^\Delta - u^\Delta) \\ &\le 4C_g(1+\sqrt{3}K_0)\Delta^{\beta/6} + 2C_g K_0 \Delta^{\beta/6} + 2C_g(1+\sqrt{3}K_0)K_1 C_\rho T \Delta^{\beta/6}, \end{aligned}$$

where we used the fact that $c(0, \Delta^{1/6}) \le C_\rho$. The conclusion then follows by combining the above estimate with (5.32). ■

5.4.3 The non-degenerate case

We now prove part (iii) in Theorem 5.1.4. When the non-degeneracy assumption and more regularity on the initial data g are imposed as in part (iii), the solution u of (5.2)-(5.3) becomes a classical solution with enough regularity. This will significantly simplify the previous proof for the general case with possible degeneracy.

First, the monotonicity property of $\hat{\mathbb{E}}$, the boundedness of g and (5.6) yield that u is bounded. Lemma 5.2.2 further implies that $u \in \mathcal{C}_b^{1/2,1}(\bar{Q}_T)$. In turn, the regularity theory of fully nonlinear PDEs implies the Hölder continuity of the derivatives of u, i.e. there exists a constant $\alpha \in (0,1)$ depending only on d, $\underline{\sigma}^2$ and M_X^2 such that $u \in \mathcal{C}_b^{1+\frac{\alpha}{2}, 2+\alpha}(\bar{Q}_T^\varepsilon)$ for any $\varepsilon > 0$ (see Theorem 4.5 in Appendix C of [54], or [41] and [66] for more details). The consistency error estimate (5.24) then yields

$$\begin{aligned} & |S(\Delta, x, u(t,x), u(t-\Delta, \cdot))| \\ \le\ & K_\alpha \left(\Delta^{\alpha/2} \left([D_x^2 u]_{\mathcal{C}^{\alpha/2,\alpha}} + [\partial_t u]_{\mathcal{C}^{\alpha/2,\alpha}} \right) + \sqrt{\Delta}|D_x^2 u|_0 + \Delta \left(|\partial_t D_x^2 u|_0 + |\partial_t D_x u|_0 \right) \right) \\ \le\ & C\Delta^{\alpha/2}, \end{aligned}$$

for $(t,x) \in Q_T^\Delta$ and some constant C. On the other hand, since

$$S(\Delta, x, u^\Delta(t,x), u^\Delta(t-\Delta, \cdot)) = 0,$$

the comparison principle in Proposition 5.3.5 implies

$$|u - u^\Delta| \le \sup_{\bar{Q}_T \setminus Q_T^\Delta} |u - u^\Delta| + Ct\Delta^{\alpha/2} \ \text{ in } \bar{Q}_T. \tag{5.41}$$

Since Assumption 5.1.2(i) holds with $\beta = 1$, it follows from Lemma 5.4.1 that

$$\sup_{\bar{Q}_T \setminus Q_T^\Delta} |u - u^\Delta| \le 2C_g K_0 \Delta^{1/2}.$$

The conclusion follows by plugging the above estimate into (5.41) and combining with part (ii) in Theorem 5.1.4.

Remark 5.4.6 *Since there is no explicit formula for the Hölder constant α,* *we are not able to write down the explicit error bound as for the general case with possible degeneracy in part(ii) of Theorem 5.1.4.*

On the other hand, if the solution u *has more regularity, say* $u \in \mathcal{C}_b^\infty(Q_T)$, *then we can replace the consistency error estimate (5.24) in the above proof by (5.26), and obtain the convergence rate* $\Delta^{1/2}$.

Chapter 6

A monotone scheme for G-equations II: the applications to sublinear expectation

This chapter continues the discussion of the monotone scheme for G-equations introduced in the last chapter but focuses mainly on some applications to sublinear expectation. A direct application of Theorem 5.1.4 is obtaining the convergence rate of Peng's robust central limit theorem with an explicit bound of Berry-Esseen type. This is an extension of the convergence rate results of classical central limit theorem (CLT) we introduced in Section 2.5. The convergence rate of robust central limit theorem we obtain improves all the existing ones obtained under different model assumptions in the literature. This is presented in Section 6.2-6.3.

Another application is an approximation scheme with its convergence rate for the Black-Scholes-Barenblatt (BSB) equation, which is widely used to model volatility uncertainty (see [1], [48] and [65]). We show that the proposed approximation scheme is a natural generalization of the well known Cox-Ross-Rubinstein (CIR) binomial tree approximation to the case with model ambiguity (see section 6.4).

Finally in Section 6.5, we make a connection between G-equations and HJB equations, and show an approximation for G-normal distribution with its convergence rate by a simple optimal switching system.

6.1 Basic notions of sublinear expectations

Here we present some basic notions of sublinear expectation framework, which is a generalization of the classical linear probability framework. Please refer to [51–54] for more details.

Let Ω be a given set and $\mathcal{H}$ be a linear space of real valued functions defined on Ω such that $X_1, X_2, ..., X_n \in \mathcal{H}$ implies that $\varphi(X_1, X_2, ..., X_n) \in \mathcal{H}$ for any $\varphi \in \mathcal{C}_{l,Lip}(\mathbb{R}^n)$, where $\varphi \in \mathcal{C}_{l,Lip}(\mathbb{R}^n)$ denotes the space of functions φ satisfying the locally Lipschitz condition

$$|\varphi(x) - \varphi(y)| \leq C(1 + |x|^k + |y|^k)|x - y|, \quad \forall x, y \in \mathbb{R}^n,$$

for some $C > 0$, $k \in \mathbb{N}$ depending on φ. $\mathcal{H}$ is considered as the space of random variables under the sublinear expectation framework.

Definition 6.1.1 *(Sublinear expectation) A sublinear expectation on $\mathcal{H}$ is a functional $\hat{\mathbb{E}} : \mathcal{H} \to \mathbb{R}$ satisfying the following axioms: for all $X, Y \in \mathcal{H}$,*

- Monotonicity*: $\hat{\mathbb{E}}[X] \geq \hat{\mathbb{E}}[Y]$ if $X \geq Y$.*
- Constant preserving*: $\hat{\mathbb{E}}[c] = c$ for $c \in \mathbb{R}$.*
- Sub-additivity*: $\hat{\mathbb{E}}[X + Y] \leq \hat{\mathbb{E}}[X] + \hat{\mathbb{E}}[Y]$.*
- Positive homogeneity*: $\hat{\mathbb{E}}[\lambda X] = \lambda \hat{\mathbb{E}}[X]$ for $\lambda \geq 0$.*

The triple $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ is called a *sublinear expectation space.* Let $X = (X_1, X_2, ..., X_n) \in \mathcal{H}^n$ be a n-dimensional random vector, and define

$$\hat{\mathbb{F}}_X[\varphi] := \hat{\mathbb{E}}[\varphi(X)] : \varphi \in \mathcal{C}_{l,Lip}(\mathbb{R}^n) \mapsto (-\infty, \infty),$$

which is a sublinear expectation on $\mathcal{C}_{l,Lip}(\mathbb{R}^n)$.

Definition 6.1.2 *(Distribution) We call $\hat{\mathbb{F}}_X$ the distribution of X. Two random vectors $X, Y \in \mathcal{H}^n$ are called identically distributed, denote as $X \overset{d}{=} Y$, if*

$$\hat{\mathbb{E}}[\varphi(X)] = \hat{\mathbb{E}}[\varphi(Y)], \quad \forall \varphi \in \mathcal{C}_{l,Lip}(\mathbb{R}^n),$$

namely, $\hat{\mathbb{F}}_X = \hat{\mathbb{F}}_Y$. A sequence of random vectors $\{X_n\}_n^\infty$ converges in distribution to X, denoted as $X_n \overset{d}{\to} X$, if

$$\lim_{n \to \infty} \hat{\mathbb{E}}[\varphi(X_n)] = \hat{\mathbb{E}}[\varphi(X)], \quad \forall \varphi \in \mathcal{C}_{l,Lip}(\mathbb{R}^n),$$

namely, $\hat{\mathbb{F}}_{X_n} \to \hat{\mathbb{F}}_X$ pointwise.

Definition 6.1.3 *(Independence) A random vector $Y \in \mathcal{H}^n$ is said to be independent to another random vector $X \in \mathcal{H}^m$, if*

$$\hat{\mathbb{E}}[\varphi(X, Y)] = \hat{\mathbb{E}}[\hat{\mathbb{E}}[\varphi(x, Y)]_{x=X}], \quad \forall \varphi \in \mathcal{C}_{l,Lip}(\mathbb{R}^{m+n}).$$

A random vector $\bar{X}$ is called an independent copy of X if $\bar{X} \overset{d}{=} X$ and $\bar{X}$ is independent to X.

Under sublinear expectations, the notion of independence is asymmetric in the sense that Y is independent to X does not imply that X is independent to Y.

Definition 6.1.4 *(Independent and identically distributed) A sequence of random vectors $\{X_n\}_n^\infty$ is said to be independent and identically distributed (i.i.d), if $X_{n+1} \overset{d}{=} X_n$ and X_{n+1} is independent to $(X_1, X_2, ..., X_n)$ for each $n \in \mathbb{N}$.*

Definition 6.1.5 *(G-distribution) A pair of random vectors (X, Y) is called G-distributed, if*

$$\big(aX + b\bar{X}, a^2Y + b^2\bar{Y}\big) \overset{d}{=} (\sqrt{a^2+b^2}X, (a^2+b^2)Y) \quad \forall a, b \in \mathbb{R},$$

where $(\bar{X}, \bar{Y})$ is an independent copy of (X, Y). Moreover, X is said to be G-normal distributed and Y is said to be maximal distributed.

Remark 6.1.6 *It can be deduce from the above definition that a G-normal random vector X has no mean uncertainty: for each component X_i of X,*

$$\hat{\mathbb{E}}[X_i] = \hat{\mathbb{E}}[-X_i] = 0, \quad i = 1, ..., d.$$

Moreover, there exists a bounded subset $\hat{\Theta} \in \mathbb{S}(d)$ such that

$$\hat{\mathbb{E}}[\langle AX, X\rangle] = \sup_{Q\in\hat{\Theta}} tr[AQ], \quad \forall A \in \mathbb{S}(d).$$

In general, $\hat{\Theta}$ characterizes the covariance uncertainty of X and the distribution of X is denoted by $X \sim \mathcal{N}(0, \hat{\Theta})$.

Similarly, for a maximal distributed random vector Y, there exists a bounded, closed and convex subset $\Gamma \subset \mathbb{R}^d$ such that

$$\hat{\mathbb{E}}[\varphi(Y)] = \max_{q\in\Gamma} \varphi(q), \quad \forall \varphi \in \mathcal{C}_{l,Lip}(\mathbb{R}^n).$$

Γ characterizes the mean uncertainty of Y and the distribution of Y is denoted by $Y \sim \mathcal{M}(\Gamma)$.

6.2 Application to robust central limit theorem

In this section, we apply Theorem 5.1.4 to derive the convergence rate (with an explicit bound of Berry-Esseen type) of the celebrated robust central limit theorem introduced in [52]. For this, let $\{(X_i, Y_i)\}_{i\geq 1}$ be a sequence of i.i.d $\mathbb{R}^d \times \mathbb{R}^d$-valued random vectors defined on $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ such that each (X_i, Y_i) is an independent copy of (X, Y). Furthermore, assume that (X, Y) satisfies Assumption 5.1.2(ii). Then, Peng proved that the sequence $\{S_n\}_{n\geq 1}$ defined by

$$S_n := \sum_{i=1}^{n} (\frac{X_i}{\sqrt{n}} + \frac{Y_i}{n}) \tag{6.1}$$

converges in distribution to $(\xi + \zeta)$:

$$\lim_{n\to\infty} \hat{\mathbb{E}}[g(S_n)] = \widetilde{\mathbb{E}}[g(\xi + \zeta)]. \tag{6.2}$$

for any continuous test function g satisfying linear growth condition. See Theorem 5.1 in [52] for its proof.

Following Peng's seminal work, a number of efforts have been made to further obtain the various convergence rates of (6.2) with additional model assumptions (see, for example, [30] [45] and [60]). However, the existing literature on the convergence rates of (6.2) assumes that either $X_i = 0$ or $Y_i = 0$ and, to be best of our knowledge, the convergence rate of (6.2) for the general situation (i.e. $X_i \neq 0$ and $Y_i \neq 0$) and an explicit bound of Berry-Esseen type are still lacking. Our aim is therefore to obtain a general result about the convergence rate of (6.2) with an explicit bound using Theorems 5.1.4.

To illustrate how it works, we provide some preliminary informal arguments to highlight the main ideas and build intuition. Consider $d = 1$ for simplicity. If we replace the sublinear expectation $\hat{\mathbb{E}}$ with the linear expectation $\mathbb{E}$ and let $\{(X_i, Y_i)\}_{i\geq 1}$ be a sequence of i.i.d. copies of (X, Y) such that $\mathbb{E}[X] = 0$, then the recursive approximation (5.7) reduces to

$$u^{\Delta}(n\Delta, x) = \mathbb{E}[g(x + \sum_{i=1}^{n}(\sqrt{\Delta}X_i + \Delta Y_i))].$$

On the other hand, the nonlinear function G defined in (5.1) reduces to $G(p, A) = \frac{1}{2}\mathbb{E}[|X|^2]A + \mathbb{E}[Y]p$, so $\mathbb{E}[|X|^2]$ and $\mathbb{E}[Y]$ turn out to be the coefficients of the linear equation

$$\partial_t u - \frac{1}{2}\mathbb{E}[|X|^2]\partial_x^2 u - \mathbb{E}[Y]\partial_x u = 0.$$

The Feynman-Kac formula then implies that

$$u(t, x) = \mathbb{E}[g(x + \sqrt{t}\xi + t\zeta)],$$

where $\xi \sim N(0, \mathbb{E}[|X|^2])$ and $\zeta = \mathbb{E}[Y]$. Taking $\Delta = \frac{1}{n}$ and using Theorem 5.1.4, we obtain

$$u^{\Delta}(1, 0) = \mathbb{E}\left[g\left(\sum_{i=1}^{n}(\frac{X_i}{\sqrt{n}} + \frac{Y_i}{n})\right)\right] \to u(1, 0) = \mathbb{E}[g(\xi + \zeta)],$$

which is precisely the classical central limit theorem (for ξ) and law of large numbers (for ζ).

Theorem 6.2.1 *Let $\{S_n\}_{n\geq 1}$ be given as in (6.1), and suppose Assumption 5.1.2 is satisfied. Then, the following assertions hold.*

(i) (Degenerate case) There exists a constant C depending only on T, C_g, β, M_X^3 and M_Y^2 such that

$$\left|\hat{\mathbb{E}}[g(S_n)] - \widetilde{\mathbb{E}}[g(\xi + \zeta)]\right| \leq Cn^{-\frac{\beta}{6}}. \tag{6.3}$$

Moreover, if the dimension $d = 1$ and $T = 1$, then the constant C has an explicit formula given in (5.9).

(ii) (Non-degenerate case) Moreover, if the second moment of the random vector X is non-degenerate, i.e.

$$\underline{\sigma}^2 := -\hat{\mathbb{E}}[-|X|^2] > 0,$$

and the initial data $g \in \mathcal{C}_b^1(\mathbb{R}^d)$, i.e. g is bounded and Lipschitz continuous, then there exists a constant $\alpha \in (0, 1)$ depending on d, $\underline{\sigma}^2$ and M_X^2 and a *constant C depending only on T, C_g, α, $M_X^{2+\alpha}$ and M_Y^2 such that*

$$\left|\hat{\mathbb{E}}[g(S_n)] - \widetilde{\mathbb{E}}[g(\xi+\zeta)]\right| \leq Cn^{-\max\{\frac{\alpha}{2},\frac{1}{6}\}}. \tag{6.4}$$

Proof. We claim that, for all $n \in \mathbb{N}$ such that $n\Delta \leq T$ and $x \in \mathbb{R}^d$,

$$u^{\Delta}(n\Delta, x) = \hat{\mathbb{E}}[g(x + \sum_{i=1}^{n}(\sqrt{\Delta}X_i + \Delta Y_i))]. \tag{6.5}$$

If the representation formula (6.5) holds, then by letting $\Delta = 1/n$ and $x = 0$, we obtain

$$u^{\Delta}(1,0) = \hat{\mathbb{E}}[g(S_n)].$$

On the other hand, the representation formula (5.6) implies that

$$u(1,0) = \widetilde{\mathbb{E}}[g(\xi+\zeta)].$$

Hence, the assertions (i) and (ii) follow from Theorem 5.1.4.

We are left to show (6.5). We prove by induction on n. Note that the case $n = 1$ follows directly from (5.7). Next, we claim that for all $n \in \mathbb{N}$ and $h \in \mathcal{C}_{lb}(\mathbb{R}^d)$,

$$\hat{\mathbb{E}}\left[h\left(\sum_{i=1}^{n}(\sqrt{\Delta}X_i + \Delta Y_i)\right)\right] = \hat{\mathbb{E}}\left[h\left(\sum_{i=2}^{n+1}(\sqrt{\Delta}X_i + \Delta Y_i)\right)\right], \tag{6.6}$$

and suppose (6.5) holds for some $n \in \mathbb{N}$ such that $n\Delta \leq T$. Then, if $(n+1)\Delta \leq T$, we use (6.6) to obtain

$$\begin{aligned}
&u^{\Delta}((n+1)\Delta, x)\\
&= \hat{\mathbb{E}}[u^{\Delta}(n\Delta, x + \sqrt{\Delta}X + \Delta Y)]\\
&= \hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[g(x + \sqrt{\Delta}p + \Delta q + \sum_{i=1}^{n}(\sqrt{\Delta}X_i + \Delta Y_i))\right]_{(p,q)=(X,Y)}\right]\\
&= \hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[g(x + \sqrt{\Delta}p + \Delta q + \sum_{i=2}^{n+1}(\sqrt{\Delta}X_i + \Delta Y_i))\right]_{(p,q)=(X,Y)}\right]\\
&= \hat{\mathbb{E}}[g(x + \sum_{i=1}^{n+1}(\sqrt{\Delta}X_i + \Delta Y_i))].
\end{aligned}$$

In other words, (6.5) also holds for $n + 1$.

Finally, to show (6.6), we prove again by induction on n. The case $n = 1$ follows from $(X_2, Y_2) \overset{d}{=} (X_1, Y_1)$. Suppose (6.6) holds for some $n \in \mathbb{N}$, then

$$\hat{\mathbb{E}}\left[h\left(\sum_{i=1}^{n+1}(\sqrt{\Delta}X_i + \Delta Y_i)\right)\right] = \hat{\mathbb{E}}\left[h\left(\sum_{i=1}^{n}(\sqrt{\Delta}X_i + \Delta Y_i) + \sqrt{\Delta}X_{n+1} + \Delta Y_{n+1}\right)\right].$$

Since (X_{n+1}, Y_{n+1}) is independent to $\{(X_1, Y_1), ..., (X_n, Y_n)\}$, The RHS of the above equality further equals to

$$
\begin{aligned}
&\hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[h\left(\sum_{i=1}^{n}(\sqrt{\Delta}x_i+\Delta y_i)+\sqrt{\Delta}X_{n+1}+\Delta Y_{n+1}\right)\right]_{(x_i,y_i)=(X_i,Y_i),i=1,\ldots,n}\right]\\
&=\hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[h\left(\sum_{i=1}^{n}(\sqrt{\Delta}x_i+\Delta y_i)+\sqrt{\Delta}X_{n+2}+\Delta Y_{n+2}\right)\right]_{(x_i,y_i)=(X_i,Y_i),i=1,\ldots,n}\right]\\
&=\hat{\mathbb{E}}\left[f\left(\sum_{i=1}^{n}(\sqrt{\Delta}X_i+\Delta Y_i)\right)\right],
\end{aligned}
$$

where $f(x):=\hat{\mathbb{E}}\left[h\left(x+\sqrt{\Delta}X_{n+2}+\Delta Y_{k+2}\right)\right]$ and the first equality follows from $(X_{n+2},Y_{n+2})\overset{d}{=}(X_{n+1},Y_{n+1})$. In turn, since (6.6) holds for n, we further have

$$
\begin{aligned}
&\hat{\mathbb{E}}\left[h\left(\sum_{i=1}^{n+1}(\sqrt{\Delta}X_i+\Delta Y_i)\right)\right]\\
&=\hat{\mathbb{E}}\left[f\left(\sum_{i=2}^{n+1}(\sqrt{\Delta}X_i+\Delta Y_i)\right)\right]\\
&=\hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[h\left(\sum_{i=2}^{n+1}(\sqrt{\Delta}x_i+\Delta y_i)+\sqrt{\Delta}X_{n+2}+\Delta Y_{n+2}\right)\right]_{(x_i,y_i)=(X_i,Y_i),i=2,\ldots,n+1}\right]\\
&=\hat{\mathbb{E}}\left[h\left(\sum_{i=2}^{n+2}(\sqrt{\Delta}X_i+\Delta Y_i)\right)\right],
\end{aligned}
$$

which completes the proof. ∎

6.3 Some special cases

In this section, we improve the convergence rates in Theorem 6.2.1 by imposing further model assumptions, and compare our results with the existing literature. For the latter use, we state the following property (see Proposition 4.1 in [52]) of the nonlinear function $G(p,A)$ given by (5.1).

Proposition 6.3.1 *Let the nonlinear function $G(p,A)$ be given in (5.1). Then, there exists a bounded, closed and compact subset $\Theta\subset\mathbb{R}^d\times\mathbb{R}^{d\times d}$ such that*

$$G(p,A)=\sup_{(q,Q)\in\Theta}\left\{\frac{1}{2}tr[AQ]+\langle p,q\rangle\right\},\ (p,A)\in\mathbb{R}^d\times\mathbb{S}(d). \tag{6.7}$$

Remark 6.3.2 *For a nonlinear function $G(p,A)$ satisfying (6.7), define $\bar{G}(p):\mathbb{R}^d\to\mathbb{R}$ by $\bar{G}(p):=G(p,0)$. The subset Θ then reduces to some $\Gamma\subset\mathbb{R}^d$, which characterizes a maximal distributed random vector $\zeta\sim\mathcal{M}(\Gamma)$.*

Similar, define $\hat{G}(A):\mathbb{S}(d)\to\mathbb{R}$ by $\hat{G}(A):=G(0,A)$. The subset Θ then reduces to some $\hat{\Theta}\subset\mathbb{S}(d)$, which characterizes a G-normal random vector $\xi\sim\mathcal{N}(0,\hat{\Theta})$.

In this sense, the nonlinear function G characterizes the above G-distributed pair (ξ, ζ) in some sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$. Moreover,

$$G(p, A) = \hat{\mathbb{E}}\left[\langle p, \zeta\rangle + \frac{1}{2}\langle A\xi, \xi\rangle\right], \quad \forall (p, A) \in \mathbb{R}^d \times \mathbb{S}(d).$$

See Proposition 4.2 of [52] for further details.

6.3.1 Law of large numbers: Comparison with [30]

Assume that $X = 0$. With this extra assumption, we can obtain a better convergence rate by refining the consistency error estimates in Proposition 5.3.1 and Proposition 5.3.3.

Corollary 6.3.3 *Suppose that Assumption 5.1.2 is satisfied with $X = 0$ and $\beta = 1$, i.e. there is no volatility uncertainty, and the initial data g is Lipschitz continuous bounded from below. Then, there exists a constant C depending only on T, Θ and M_Y^2 such that*

$$|u - u^{\Delta}| \le C\Delta^{\frac{1}{2}} \text{ in } \bar{Q}_T.$$

Before proving Corollary 6.3.3, we show its application to the generalized law of large numbers. To this end, let $\Delta = \frac{1}{n}$, then by the representation formula (6.5), we have

$$u^{1/n}(1, 0) = \hat{\mathbb{E}}[g(\sum_{i=1}^{n} \frac{Y_i}{n})],$$

where $Y_1 = Y$, $Y_{i+1} \overset{d}{=} Y_i$ and Y_{i+1} is independent to $(Y_1, ..., Y_i)$ for each $i = 1, ..., n-1$. On the other hand, if we further let $g(y) := d_\Gamma(y) = \inf\{|x - y| : x \in \Gamma\}$, where the subset $\Gamma \subset \mathbb{R}^d$ is given Remark 6.3.2, then $d_\Gamma(y) \ge 0$ is Lipschitz continuous bounded from below. It follows from (5.6) and Remark 6.1.6 that

$$u(1, 0) = \sup_{\theta \in \Gamma} g(\theta) = \sup_{\theta \in \Gamma} d_\Gamma(\theta) = 0.$$

In turn, Corollary 6.3.3 yields the following form of *generalized law of large numbers*

$$0 \le \hat{\mathbb{E}}[d_\Gamma(\sum_{i=1}^{n} \frac{Y_i}{n})] \le Cn^{-1/2}. \tag{6.8}$$

Note that the above convergence rate is better than the convergence rate $n^{-2/5}$ in Fang et al [30] for the law of large numbers under sublinear expectations (See Remark 2.3 in [30]).

Remark 6.3.4 *If we choose any bounded and Lipschitz continuous $g \in \mathcal{C}_b^1(\mathbb{R}^d)$ as the test function (which clearly satisfies the assumption in Corollary 6.3.3), then we also obtain the following general form of generalized law of large numbers*

$$\left|\hat{\mathbb{E}}[g(\sum_{i=1}^{n} \frac{Y_i}{n})] - \sup_{\theta \in \Gamma} g(\theta)\right| \le Cn^{-1/2}. \tag{6.9}$$

We proceed to prove Corollary 6.3.3. Since it is a special case of Theorem 5.1.4, we only highlight its main steps and differences compared to the proof of Theorem 5.1.4. Unless otherwise specified, C will represent a generic constant in the following.

Step 1. Since $X = 0$ and $\beta = 1$, a revisit of Lemmas 5.2.2 and 5.2.3 shows that u and u^Δ satisfy

$$|u(t,x) - u(s,y)| \le C_g[|x-y| + M_Y^1|s-t|];$$
$$|u^\Delta(t,x) - u^\Delta(s,y)| \le C_g[|x-y| + M_Y^1(|s-t| + \Delta)].$$

Thus, from Lemma 5.4.1, the error bound between u and u^Δ in the interval $[0,\Delta]$ can be refined as

$$\sup_{\bar{Q}_T \setminus Q_T^\Delta} |u - u^\Delta| \le C\Delta. \tag{6.10}$$

Step 2. Next, we refine the consistency error estimates. From Proposition 5.3.1, since $X = 0$, term (I) disappears and $(II) \le \frac{1}{2}\Delta^2|D^2\psi|_0 M_Y^2$. Thus, $\mathcal{E}(\Delta,\psi) \le C\Delta|D^2\psi|_0$. Plugging it into Proposition 5.3.3(iii) yields

$$|\partial_t\psi - G(D_x\psi,0) - S(\Delta,x,\psi,\psi(t-\Delta,\cdot))| \le C\Delta\left(|D_x^2\psi|_0 + |\partial_t^2\psi|_0 + |\partial_t D_x\psi|_0\right). \tag{6.11}$$

Step 3. We modify the mollifiers ρ_ε in Section 5.4.1 by

$$\rho_\varepsilon(t,x) := \frac{1}{\varepsilon^{1+d}}\rho\left(\frac{t}{\varepsilon},\frac{x}{\varepsilon}\right), \tag{6.12}$$

and redefine u_ε as

$$u_\varepsilon(t,x) = u * \rho_\varepsilon(t,x) = \int_{-\varepsilon<\tau<0}\int_{|e|<\varepsilon} u(t-\tau, x-e)\rho_\varepsilon(\tau,e)\,de\,d\tau.$$

Thus, the regularity of u implies that $|u - u^\varepsilon| \le C\varepsilon$, and $|\partial_t^i D_x^j u_\varepsilon|_0 \le C\varepsilon^{1-i-|j|}$.

Step 4. Substituting u_ε with ψ in the consistency error estimate (6.11) and using the fact that $\partial_t u_\varepsilon - G(D_x u_\varepsilon, 0) \ge 0$ yield

$$S(\Delta, x, u_\varepsilon(t,x), u_\varepsilon(t-\Delta,\cdot)) \ge -C\Delta\varepsilon^{-1}.$$

Furthermore, choosing $\varepsilon = \Delta^{1/2}$ and following along the similar arguments as in the proof for Theorem 5.4.3, we obtain that

$$u^\Delta - u \le \sup_{\bar{Q}_T \setminus Q_{T-\Delta}} (u - u^\Delta)^+ + C\Delta^{1/2} \le C\Delta^{1/2},$$

where we used (6.10) in the last inequality.

Step 5. To prove the other side inequality, we mollify u^Δ with $\rho_\varepsilon(t,x)$ given in (6.12), i.e. $u_\varepsilon^\Delta(t,x) = u^\Delta * \rho_\varepsilon(t,x)$. Then, the regularity of u^Δ implies that

$$|u^\Delta - u_\varepsilon^\Delta|_0 \le C(\varepsilon + \Delta),$$

and

$$|\partial_t^i D_x^j u_\varepsilon^\Delta|_0 \le C\varepsilon^{-i-|j|}(\varepsilon + \Delta),$$

Step 6 Substituting u_ε^Δ with ψ in the consistency error estimate (6.11) and using the fact that $S(\Delta, x, u_\varepsilon^\Delta(t,x), u_\varepsilon^\Delta(t-\Delta,\cdot)) \geq 0$ yield

$$\partial_t u_\varepsilon^\Delta - G(D_x u_\varepsilon^\Delta, 0) \geq -C(\varepsilon + \Delta)\Delta\varepsilon^{-2},$$

In turn, $u - u_\varepsilon^\Delta \leq C(\varepsilon+\Delta)(1+\Delta\varepsilon^{-2})$, and by choosing $\varepsilon = \Delta^{1/2}$ and following along the similar arguments as in the proof for Theorem 5.4.5, we obtain

$$u - u^\Delta = (u - u_\varepsilon^\Delta) + (u_\varepsilon^\Delta - u^\Delta) \leq C\Delta^{1/2},$$

which is the desired convergence rate.

6.3.2 Central limit theorem: Comparison with [45] and [60]

Assume that $Y = 0$, we obtain the central limit theorem as in [45] and [60], but with an improved convergence rate. To this end, choosing $\Delta = 1/n$, by the representation formula (6.5), we have

$$u^{1/n}(1,0) = \hat{\mathbb{E}}[g(\sum_{i=1}^{n} \frac{X_i}{\sqrt{n}})],$$

where $X_1 = X$, $X_{i+1} \overset{d}{=} X_i$ and X_{i+1} is independent to $(X_1, ..., X_i)$ for each $i = 1, ..., n-1$. On the other hand, since $G(p, A) = \hat{\mathbb{E}}[\frac{1}{2}\langle AX, X\rangle]$ and $\zeta = 0$, by (5.6), we have

$$u(1,0) = \widetilde{\mathbb{E}}[g(\xi)] = \mathcal{N}_G(g),$$

where $\mathcal{N}_G$ denotes the corresponding G-normal distribution. Under Assumption 5.1.2, Theorem 6.2.1 then yields the following *robust central limit theorem* in the degenerate case

$$\left| \hat{\mathbb{E}}[g(\sum_{i=1}^{n} \frac{X_i}{\sqrt{n}})] - \mathcal{N}_G(g) \right| \leq Cn^{-\beta/6}. \tag{6.13}$$

Moreover, if the second moment of the random vector X is non-degenerate, i.e. $\underline{\sigma}^2 := -\hat{\mathbb{E}}[-|X|^2] > 0$, and the initial data $g \in \mathcal{C}_b^1(\mathbb{R}^d)$, i.e. g is bounded and Lipschitz continuous, then

$$\left| \hat{\mathbb{E}}[g(\sum_{i=1}^{n} \frac{X_i}{\sqrt{n}})] - \mathcal{N}_G(g) \right| \leq Cn^{-\max\{\frac{\alpha}{2}, \frac{1}{6}\}}. \tag{6.14}$$

Note that the above convergence rate in (6.13) for the degenerate case improves Theorem 1.1 of Krylov [45], where the author considers a one-dimensional stochastic control problem and obtains the convergence rate $\frac{\beta^2}{4+2\beta} (\leq \frac{\beta}{6})$. This is because instead of assuming $M_X^3 < \infty$, Krylov [45] assumes only $M_X^{2+\beta} < \infty$ and thus he could only apply the consistency error estimate (5.24) (rather than (5.26)) with α replaced by β. Moreover, the convergence rate in (6.14) for the non-degeneate case improves Theorem 4.5 of Song [60], where the author obtains the convergence rate $\frac{\alpha}{2}$.

6.4 Application to Black-Scholes-Barenblatt equation

In this section, we apply the approximation scheme (5.7) to the Black-Scholes-Barenblatt equation (see [1], [48] and [65] for the dimension $d = 1$), which often arises from option pricing models with volatility uncertainty, namely

$$\partial_t u + rx\partial_x u + \frac{1}{2}\bar{\sigma}^2 x^2 \partial_{xx} u^+ - \frac{1}{2}\underline{\sigma}^2 x^2 \partial_{xx} u^- - ru = 0, \quad u|_{t=T} = g, \tag{6.15}$$

where r is the constant riskless interest rate, $\bar{\sigma} \geq \underline{\sigma} > 0$ are two constants representing upper and lower bounds on the volatility of underlying price, and $g : \mathbb{R} \to \mathbb{R}$ represents some European contingent claim payoff function. Note that when $\bar{\sigma} = \underline{\sigma}$, the equation reduces to the classical Black-Scholes equation.

To apply the approximation scheme (5.7), some transformations are needed firstly: let $v(t, x) := u(T - t, e^x)e^{rt}$, then (6.15) becomes

$$\partial_t v - \sup_{\sigma \in [\underline{\sigma}, \bar{\sigma}]} \left\{ (r - \frac{1}{2}\sigma^2)\partial_x v + \frac{1}{2}\sigma^2 \partial_{xx} v \right\} = 0, \quad v|_{t=0} = g(e^x). \tag{6.16}$$

Comparing the equation (6.16) to the G-equation (5.1) and (5.2), we only need to construct a sublinear expectation $\hat{\mathbb{E}}$ and find random variables (X, Y) with X having no mean uncertainty such that

$$G(p, A) = \hat{\mathbb{E}}[pY + \frac{1}{2}AX^2] = \sup_{\sigma \in [\underline{\sigma}, \bar{\sigma}]} \left\{ (r - \frac{1}{2}\sigma^2)p + \frac{1}{2}\sigma^2 A \right\}. \tag{6.17}$$

To this end, suppose we are given a measurable space $(\Omega, \mathcal{F})$ which supports a class of probability measures $\mathbb{P}^\sigma$ for $\sigma \in [\underline{\sigma}, \bar{\sigma}]$. We can then define a random variable X such that $\mathbb{P}^\sigma(X = \sigma) = \frac{1}{2}$ and $\mathbb{P}^\sigma(X = -\sigma) = \frac{1}{2}$ for any $\sigma \in [\underline{\sigma}, \bar{\sigma}]$, and a random variable $Y := r - \frac{1}{2}X^2$. Consequently, a sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ can be defined such that $X, Y \in \mathcal{H}$ and

$$\hat{\mathbb{E}}[\xi] = \sup_{\sigma \in [\underline{\sigma}, \bar{\sigma}]} \mathbb{E}^{\mathbb{P}^\sigma}[\xi] \quad \text{for} \quad \xi \in \mathcal{H}.$$

It is clear that $\hat{\mathbb{E}}[X] = \hat{\mathbb{E}}[-X] = 0$ and (6.17) holds. The approximation scheme (5.7) then has a simple form:

$$\begin{aligned} v^\Delta(t, x) &= \hat{\mathbb{E}}[v^\Delta(t - \Delta, x + \sqrt{\Delta}X + \Delta Y)] \\ &= \sup_{\sigma \in [\underline{\sigma}, \bar{\sigma}]} \mathbb{E}^{\mathbb{P}^\sigma}[v^\Delta(t - \Delta, x + \sqrt{\Delta}X + \Delta(r - \frac{1}{2}X^2))] \\ &= \sup_{\sigma \in [\underline{\sigma}, \bar{\sigma}]} \left[\frac{1}{2}v^\Delta \left(t - \Delta, x + (r - \frac{1}{2}\sigma^2)\Delta + \sigma\sqrt{\Delta} \right) \right. \\ &+ \left. \frac{1}{2}v^\Delta \left(t - \Delta, x + (r - \frac{1}{2}\sigma^2)\Delta - \sigma\sqrt{\Delta} \right) \right] \end{aligned} \tag{6.18}$$

for $\Delta \leq t \leq T$ and $v^\Delta(t, x) = g(e^x)$ for $t < \Delta$.

Remark 6.4.1 *When $\bar{\sigma} = \underline{\sigma} = \sigma > 0$, the above approximation scheme reduces to classical Cox, Ross and Rubinstein (CRR) binomial tree approximation* *for $X_t = \ln(S_t)$, with S_t following the geometric Brownian motion $dS_t = rS_t dt + \sigma S_t dW_t$. Indeed, by Itô's lemma, we have $dX_t = (r - \frac{1}{2}\sigma^2)dt + \sigma dW_t$. The CRR binomial tree approximation for X_t is then given as follows (we only display one step binomial tree approximation for simplicity):*

$$
\begin{tikzcd}
 & X_0 + (r - \frac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{\Delta t} \\
X_0 \arrow[ur, "\frac{1}{2}"] \arrow[dr, "\frac{1}{2}"'] & \\
 & X_0 + (r - \frac{1}{2}\sigma^2)\Delta t - \sigma\sqrt{\Delta t}
\end{tikzcd}
$$

The approximation scheme for the approximation of $v(t,x)$ then reduces to the CRR binomial tree approximation

$$
\begin{aligned}
v^\Delta(t,x) = &\ \frac{1}{2}v^\Delta\left(t - \Delta, x + (r - \frac{1}{2}\sigma^2)\Delta + \sigma\sqrt{\Delta}\right) \\
&+ \frac{1}{2}v^\Delta\left(t - \Delta, x + (r - \frac{1}{2}\sigma^2)\Delta - \sigma\sqrt{\Delta}\right).
\end{aligned}
$$

Since Assumption 5.1.2(ii) clearly holds, if the composition function $\phi(x) := g(e^x)$ satisfies Assumption 5.1.2(i), Theorem 5.1.4 then implies that v^Δ converges to v locally uniformly. Notice that with our construction of the sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ and our choice of the random variable X, it also holds that

$$
\hat{\mathbb{E}}[X] = \hat{\mathbb{E}}[-X] = \hat{\mathbb{E}}[X^3] = \hat{\mathbb{E}}[-X^3] = 0, \quad M_X^4 = \hat{\mathbb{E}}[|X|^4] < \infty. \tag{6.19}
$$

Thanks to the above properties, we obtain a better convergence rate than that in Theorem 5.1.4(iii).

Proposition 6.4.2 *Let (X, Y) be the random variables constructed as above on the sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$. Then, for any test function $\phi(x) := g(e^x) \in \mathcal{C}_b^1(\mathbb{R})$, there exists a constant $\alpha \in (0,1)$ depending on $\underline{\sigma}^2$ and M_X^2 such that $v \in \mathcal{C}_b^{1+\frac{\alpha}{2}, 2+\alpha}(Q_T)$. Moreover, for $\Delta \in (0,1)$, there exists a constant C depending only on T, C_g, α, and $M_X^{2+\alpha}$ such that*

$$
|v - v^\Delta| \le C\Delta^{\max\{\frac{\alpha}{2}, \frac{1}{4}\}} \quad \text{in } \bar{Q}_T. \tag{6.20}
$$

The proof follows along a similar argument and procedure used in Theorem 5.1.4 with a refinement of the consistency error estimates in Proposition 5.3.1, and is therefore postponed to the Appendix C.

Remark 6.4.3 *In the general degenerate situation (i.e. without the assumption that $\underline{\sigma} > 0$), the convergence rate in (6.20) becomes $\frac{1}{4}$. Note that (6.19) is the only condition on X needed to obtain the convergence rate $\frac{1}{4}$. The condition (6.19) is also imposed in [45] and the same convergence rate is obtained in Theorem 4.1 therein. However, there is no Y component in [45], so the result therein is only a special case of our situation.*

Remark 6.4.4 *The scheme (6.18) belongs to the class of semi-Lagrangian schemes as in [21, 29, 57], and specifically is a special case of the schemes proposed by Picarelli and Reisinger in [57] with $M = 2$, where M denotes the number of sample points used in Gauss–Hermite quadrature approximation therein. The achieved convergence rate of $\frac{1}{4}$ in the degenerate case in (6.20) is identical to that in [57] with $M = 2$, although they obtain a better convergence rate for one-sided lower bound when $M > 2$. Moreover, the scheme (6.18) is a non-recombining tree where the number of nodes of all trajectories grows exponentially in the number of time steps. In order to be able to implement the numerical computation within reasonable complexity, we typically introduce a linear interpolation type of projection step onto some space grid to provide recombination. See [56, 57] for more details.*

6.5 Optimal switching approximation for G-normal distribution

In this section, we make a connection between G-equations and HJB equations, and notice that in the one dimensional case, the control set consists of only two elements (see equation (6.26)). This inspires one way to approximate G-normal distribution by a simple two-state optimal switching system. We then give its convergence result in Theorem 6.5.3.

Let ξ be a G-normal distributed random variable in some sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ characterized by the following G-heat equation defined on $(0, 1] \times \mathbb{R}^d$:

$$\partial_t u - G(D_x^2 u) = 0, \tag{6.21}$$

with initial condition

$$u|_{t=0} = g. \tag{6.22}$$

where, from Proposition 6.3.1 and Remark 6.3.2, we have

$$G(A) = \hat{\mathbb{E}}\left[\frac{1}{2}\langle A\xi, \xi\rangle\right] = \sup_{Q\in\hat{\Theta}}\left\{\frac{1}{2}\mathrm{tr}[AQ]\right\}, \quad \forall A \in \mathbb{S}(d),$$

for some $\hat{\Theta} \subset \mathbb{S}(d)$. Thus, the G-heat equation (6.21) can be write as an HJB equation, namely,

$$\partial_t u - \sup_{Q\in\hat{\Theta}}\left\{\frac{1}{2}\mathrm{tr}\left(QD_x^2 u\right)\right\} = 0. \tag{6.23}$$

On the other hand, by (5.6), we have that (6.21)-(6.22) admits a unique viscosity solution u which admits the representation

$$u(t, x) = \hat{\mathbb{E}}[g(x + \sqrt{t}\xi)], \tag{6.24}$$

provided that the initial data g satisfies some regularity condition. The corresponding G-normal distribution is then given by

$$\mathcal{N}_G(g) = \hat{\mathbb{E}}[g(\xi)] = u(1, 0).$$

Notice that in the one dimensional case $d = 1$,

$$G(A) = \frac{1}{2}\hat{\mathbb{E}}[A\xi^2] = \frac{1}{2}(\bar{\sigma}^2 A^+ - \underline{\sigma}^2 A^-) = \frac{1}{2}\max\{\bar{\sigma}^2 A, \underline{\sigma}^2 A\}, \tag{6.25}$$

where $\bar{\sigma}^2 = \hat{\mathbb{E}}[\xi^2], \underline{\sigma}^2 = -\hat{\mathbb{E}}[-\xi^2] \geq 0$. We assume $\bar{\sigma} > \underline{\sigma}$ since otherwise (6.21) reduces to standard heat equation. Equation (6.23) then reduces to

$$\partial_t u - \max_{\sigma \in \{\bar{\sigma}, \underline{\sigma}\}} \left\{ \frac{1}{2}\sigma^2 D_x^2 u \right\} = 0. \tag{6.26}$$

Since there are only two possible values $(\bar{\sigma}, \underline{\sigma})$ for the control variable σ, it is natural to approximate the viscosity u by a two-state linear optimal switching stochastic control system on $(0, 1] \times \mathbb{R}$

$$\begin{cases} \min \left\{ \partial_t v_i - \frac{1}{2} Q_i^2 \partial_{xx} v_i, v_i - \mathcal{M}_i^k v \right\} = 0, \\ v_i|_{t=0} = g, \end{cases} \tag{6.27}$$

for $i \in \mathcal{I} := \{1, 2\}$, $(Q_1, Q_2) = (\bar{\sigma}, \underline{\sigma})$, and $\mathcal{M}_i^k v := \max_{j \neq i,\, j \in \mathcal{I}}\{v_j - k\}$, for some constant $k > 0$ representing the switching cost.

Under some regularity condition on g, the viscosity solution of optimal switching system (6.27) can then be represented on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ by

$$v_i(t, x) = \sup_{\theta \in \Theta_i[0,t]} \mathbb{E}\left[g(X_t^\theta) - \Sigma_{j\geq 1} k \mathbf{1}_{\{0 \leq \tau_j \leq t\}} \right],$$

with the state variable

$$X_t^\theta = x + \int_0^t Q_{\theta_s} dW_s,$$

where $\Theta_i[0, t]$ is the space of all admissible continuous switching control processes on $[0, t]$ starting from state i, in which any specific control process can be identified as

$$\theta_s = \Sigma_{n\geq 1} \zeta_{n-1} \mathbf{1}_{[\tau_{n-1}, \tau_n)}(s),$$

with $\{\tau_n\}_{n\geq 0}$ being a sequence of nondecreasing stopping time representing the switching time with

$$0 = \tau_0 \leq \tau_1 \leq \tau_2 \leq \cdots, \quad \text{a.s.}$$

and each ζ_n being a $\mathcal{F}_{\tau_n}$-measurable random variable valued in $\mathcal{I} = \{1, 2\}$ representing the new state after switching at τ_n, with $\zeta_0 = i$ and $\zeta_n \neq \zeta_{n+1}$ a.s.

In fact, both v_1 and v_2 converge uniformly to u as $k \to 0$. To prove this convergence and determine its convergence rate, we impose the following regularity condition on the initial condition g:

Assumption 6.5.1 *The initial data g is bounded and Lipschitz continuous with $|g|_1 \leq M$, for some constant $M > 0$.*

Assumption 6.5.1 ensures the following standard existence, uniqueness and regularity result for the switching system (6.27).

Proposition 6.5.2 *Suppose that Assumption 6.5.1 is satisfied, then*

(i) There exists a unique bounded and Lipschitz continuous solution $v = (v_1, v_2)$ of the optimal switching system (6.27) such that $|v|_1 \le C$ for some constant C depending only on M.

(ii) Standard comparison result holds, i.e. if $\underline{v}$ and $\bar{v}$ are subsolution and supersolution of (6.27) respectively, then $\underline{v}_i \le \bar{v}_i$ for each $i \in \mathcal{I}$.

Proof. See Proposition 2.1 in [4]. ■

We now prove the following error estimate between u and v_i, which gives the convergence and convergence rate of optimal switching approximation for G-normal distribution.

Theorem 6.5.3 *Suppose that Assumption 6.5.1 is satisfied. Then, for $i \in \mathcal{I} = \{1, 2\}$ we have*

$$0 \le u - v_i \le Ck^{\frac{1}{3}} \text{ in } [0,1] \times \mathbb{R},$$

for some constant C depending only on $\bar{\sigma}$ and M.

Proof. We first prove the first inequality. It can be checked that $w = (u, u)$ is a supersolution of (6.27), and then the first inequality follows directly from the comparison result Proposition 6.5.2(ii).

To prove the second inequality, we regularize each v_i by

$$v_{i,\varepsilon}(t,x) = v_i * \rho_\varepsilon(t,x) = \int_{-\varepsilon^2<\tau<0} \int_{|e|<\varepsilon} v_i(t-\tau, x-e)\rho_\varepsilon(\tau, e)ded\tau.$$

By Proposition 6.5.2(i), we know v_i is Lipschitz continuous, then standard properties of mollifiers imply that

$$|v_i - v_{i,\varepsilon}|_0 \le C\varepsilon, \tag{6.28}$$

for some constant C depending only on M.

Next, from the equations (6.27), we know that (i) $\partial_t v_i - \frac{1}{2}Q_i^2 \partial_{xx} v_i \ge 0$, and (ii) $|v_1 - v_2| \le k$. From (i), we deduce using the stability of viscosity solutions (see Proposition 2.2.6) that

$$\partial_t v_{i,\varepsilon} - \frac{1}{2}Q_i^2 \partial_{xx} v_{i,\varepsilon} \ge 0. \tag{6.29}$$

From (ii), we deduce using standard properties of mollifiers that

$$|\partial_t v_{1,\varepsilon} - \partial_t v_{2,\varepsilon}| \le Ck\varepsilon^{-2}; \quad |\partial_{xx} v_{1,\varepsilon} - \partial_{xx} v_{2,\varepsilon}| \le Ck\varepsilon^{-2},$$

for some constant C depending only on the choice of the mollifier ρ. From these estimates, we have

$$\left| \partial_t v_{j,\varepsilon} - \frac{1}{2}Q_j^2 \partial_{xx} v_{j,\varepsilon} - \partial_t v_{i,\varepsilon} + \frac{1}{2}Q_j^2 \partial_{xx} v_{i,\varepsilon} \right| \le Ck\varepsilon^{-2},$$

for $i \ne j \in \mathcal{I}$ and some constant C depending only on $\bar{\sigma}$. This together with

(6.29) when substituting i by j gives

$$\partial_t v_{i,\varepsilon} - \frac{1}{2}Q_j^2 \partial_{xx} v_{i,\varepsilon} \geq -Ck\varepsilon^{-2}.$$

Combining this with (6.29) again yields

$$\partial_t v_{i,\varepsilon} - \max_{\sigma \in \{\bar{\sigma}, \underline{\sigma}\}} \left\{ \frac{1}{2}\sigma^2 D_x^2 v_{i,\varepsilon} \right\} \geq -Ck\varepsilon^{-2}.$$

This means $\bar{v} := v_{i,\varepsilon} + Ckt\varepsilon^{-2}$ is a supersolution of (6.26) with $\bar{v}(0,x) = v_{i,\varepsilon}(0,x)$. On the other hand, from (6.28) we have $v_{i,\varepsilon}(0,x) \geq v_i(0,x) - C\varepsilon = u(0,x) - C\varepsilon$, which makes $\underline{v} := u - C\varepsilon$ a (sub)solution of (6.26) with $\underline{v}(0,x) \leq \bar{v}(0,x)$. Thus, standard comparison result for HJB equation implies that $\underline{v} \leq \bar{v}$ uniformly, i.e.

$$u \leq v_{i,\varepsilon} + C\varepsilon + Ckt\varepsilon^{-2}.$$

Using again the estimate (6.28), we finally have

$$u - v_i \leq v_{i,\varepsilon} - v_i + C\varepsilon + Ckt\varepsilon^{-2} \leq C\varepsilon + Ckt\varepsilon^{-2} = Ck^{\frac{1}{3}},$$

by choosing $\varepsilon = k^{\frac{1}{3}}$. This proves the second inequality and finishes the proof.
■

Chapter 7

Summary and Future Work

7.1 Summary

This thesis was mainly concerned with monotone schemes for non-linear second-order parabolic partial differential equations arising in applied mathematics, and their convergence analysis. Unlike the traditional numerical schemes, such as finite difference method, which discretize both time interval and space domain, the proposed schemes involve only semi-discretization in time. For each proposed approximation scheme, we proved its convergence (to viscosity solution) and establish the convergence rate by deriving the error bound (between the scheme solution and the equation viscosity solution) in the form of some exponent of the time step scaled by a constant. The exponent is then considered naturally as the convergence rate.

To do so, the thesis began by reviewing the basic theory of viscosity solutions and convergence analysis in the case of linear equations. This led to the application to the convergence rate of classical Central Limit Theorem. The first main part of the thesis then considered a specific type of semi-linear equations:

- In Chapter 3, we proposed an approximation scheme for a class of semi-linear parabolic equations (see (3.1)) whose Hamiltonian is convex and coercive to the gradients. The scheme is based on splitting the equation in two parts, the first corresponding to a linear parabolic equation and the second to a Hamilton-Jacobi equation. The solutions of theses equations are approximated using, respectively, the Feynman-Kac and the Hopf-Lax formula. We established the convergence of the approximation scheme and determined the convergence rate, combining Krylov's shaking coefficients technique and the Barles-Jakobsen optimal switching approximation. One of the key steps is the derivation of a consistency error via convex duality arguments, using the convexity of the Hamiltonian in an essential way.

- In Chapter 4, we considered the variational inequality version of equation (3.1) by adding an obstacle term f, and proposed the adapted approximation scheme based on the splitting approximation scheme in Chapter 3. The approximation error bounds were obtained similarly by Krylov's shaking coefficients technique and Barles-Jakobsen's optimal switching approximation, while the switching system used in this case is an obstacle switching system introduced in Section 4.3. The convergence rate of the scheme, however, stays the same as that in Chapter 3.

The second part of the thesis considered a class of fully nonlinear equations called G-equations, and introduced some applications in the sublinear expectation framework:

- In Chapter 5, we built a piece-wise constant approximation scheme for G-equations, and determined its convergence rate with an explicit error bound between the approximate solution and the viscosity solution of the G-equation. We obtained the error bound on one side by a standard mollification procedure, and a symmetric error bound on the other side by interchanging the roles of equation and scheme.

- Chapter 6 introduced two applications of the theoretical convergence result (Theorem 5.1.4) in Chapter 5. The first one was the derivation of the convergence rate for the *robust central limit theorem*. Thanks to the explicit formula for the error bound C in (5.9), we are able to obtain *an explicit convergence rate of Berry-Esseen type*. To the best of our knowledge, this is the first result about Peng's robust central limit theorem with an explicit bound. The second application was obtaining a numerical approximation scheme for the Black-Scholes-Barenblatt (BSB) equation, which is widely used to model volatility uncertainty. We made a connection between the G-equation and the BSB equation, and showed that the proposed approximation scheme is a natural generalization of the well known Cox-Ross-Rubinstein (CRR) binomial tree approximation to the case with model ambiguity. Finally, we presented an optimal switching approximation to G-normal distribution with its convergence rate.

7.2 Future Work

The approach and the results in the thesis may be extended in various directions. Firstly, one may consider problem (3.1) in a bounded domain, an undoubtedly important case since various applications are cast in bounded space domains (e.g. utilities defined in half-space, constrained risk measures, etc). However, various non-trivial technical difficulties arise. Some recent works on such problems using other approaches are [12], [42] and [58]. Second, one may also consider another version of variational inequalities of (3.1) where the gradient of the solution is constrained rather than the solution itself. These are naturally related to singular stochastic optimization problems. An early result in this direction but in elliptic case can be found in [25].

Another direction of future work is to relax the assumptions we posed. We mention that the approaches and results in Chapter 3 and 4 rely heavily on the Lipschitz continuity of (viscosity) solutions of the equation (3.1) and (4.1) with respect to the space variable x. A possible extension is to relax the assumptions such that the solutions are β-Hölder continuous for some $\beta \in (0, 1)$. This is challenging in the sense that in this case, (3.1) can not be written as (3.34) where the control set K is compact, which is a key step when obtaining the lower bound.

There are also several directions of future work on the second part of the thesis. Regarding to the explicit convergence rate of robust central limit theorem (see Theorem 6.2.1), the explicit constant C (given in (5.9)) is not sharp, and it would be interesting to improve such an explicit bound in the future. Also, when we applied the approach and results in the thesis to establish a Berry-Esseen type of convergence for classical central limit theorem in Section 2.5.1, the convergence rate obtained is $1/8$ (see (2.49)), which is slower than the rate of $1/2$ (see (2.47)) in the literature. Trying to recover the same convergence rate of $1/2$ could be left as future work. Finally, in Section 6.5, the optimal switching approximation for G-normal distribution was only considered in the one dimensional case. An extension to multi-dimensional case could also be potential work.

Appendix A

Proofs of Proposition 3.2.2 and 3.3.3

We note that equation (3.1) is a special case (choosing $\varepsilon = 0$) of the HJB equation (3.28). Therefore, we omit the proof of Proposition 3.2.2 and only prove Proposition 3.3.3.

We first show that there exists a bounded solution to (3.28). To this end, using the convex dual function $L^\theta(t,x,q) := \sup_{p\in\mathbb{R}^d}(p\cdot q - H^\theta(t,x,p))$, we rewrite (3.28) as

$$\begin{cases} -\partial_t u^\varepsilon + \sup\limits_{\theta\in\Theta^\varepsilon, q\in\mathbb{R}^d} \mathcal{L}^{\theta,q}\left(t,x,D_x u^\varepsilon, D_x^2 u^\varepsilon\right) = 0 & \text{in } Q_{T+\varepsilon^2}; \\ u^\varepsilon(T+\varepsilon^2, x) = U(x) & \text{in } \mathbb{R}^d, \end{cases} \tag{A.1}$$

where

$$\mathcal{L}^{\theta,q}(t,x,p,X) = -\frac{1}{2}\mathrm{tr}\left(\sigma^\theta \sigma^{\theta^T}(t,x)X\right) - (b^\theta(t,x) - q)\cdot p - L^\theta(t,x,q).$$

We also introduce the stochastic control problem

$$u^\varepsilon(t,x) = \inf_{\theta\in\Theta^\varepsilon[t,T+\varepsilon^2], q\in\mathbb{H}^2[t,T+\varepsilon^2]} \mathbb{E}\left[\int_t^{T+\varepsilon^2} L^{\theta_s}\left(s, X_s^{t,x;\theta,q}, q_s\right) ds + U\left(X_{T+\varepsilon^2}^{t,x;\theta,q}\right) |\mathcal{F}_t\right],$$

with the controlled state equation

$$dX_s^{t,x;\theta,q} = \left(b^{\theta_s}(s, X_s^{t,x;\theta,q}) - q_s\right) ds + \sigma^{\theta_s}\left(s, X_s^{t,x;\theta,q}\right) dW_s,$$

where $\Theta^\varepsilon[t,T+\varepsilon^2]$ is the space of Θ^ε-valued progressively measurable processes (τ_s, e_s) and $\mathbb{H}^2[t,T+\varepsilon^2]$ is the space of square-integrable progressively measurable processes q_s, for $s\in[t,T+\varepsilon^2]$. Next, we identify its value function with a bounded viscosity solution to (A.1). For this, we only need to establish upper and lower bounds for the value function $u^\varepsilon(t,x)$ and, in turn, use standard arguments as in [55] and [64].

To find an upper bound for u^ε, we choose an arbitrary perturbation parameter process $\theta \in \Theta^\varepsilon[t,T+\varepsilon]$ and just put $\hat{q}$ with $\hat{q}_s \equiv 0$. Then, Proposition

3.2.3 (ii) yields

$$u^\varepsilon(t,x) \leq \mathbb{E}\left[\int_t^{T+\varepsilon^2} L^{\theta_s}(s, X_s^{t,x;\theta,\hat{q}}, 0)ds + U(X_{T+\varepsilon^2}^{t,x;\theta,\hat{q}})|\mathcal{F}_t\right]$$
$$\leq (T+\varepsilon^2 - t)|L^*(0)| + M \leq (T+1)|L^*(0)| + M.$$

For the lower bound, we use again Proposition 3.2.3 (ii) to obtain that $L_*(q) \geq -H^*(0) \geq -|H^*(0)|$, for any $q \in \mathbb{R}^d$. In turn, for any $(\theta, q) \in \Theta^\varepsilon[t, T+\varepsilon^2] \times \mathbb{H}^2[t, T+\varepsilon^2]$,

$$\begin{aligned}
&\mathbb{E}\left[\int_t^{T+\varepsilon^2} L^{\theta_s}(s, X_s^{t,x;\theta,q}, q_s)ds + U(X_{T+\varepsilon^2}^{t,x;\theta,q})|\mathcal{F}_t\right] \\
&\geq \mathbb{E}\left[\int_t^{T+\varepsilon^2} L_*(q_s)ds + U(X_{T+\varepsilon^2}^{t,x;\theta,q})|\mathcal{F}_t\right] \\
&\geq -(T+\varepsilon^2 - t)|H^*(0)| - M \geq -(T+1)|H^*(0)| - M,
\end{aligned}$$

and, thus, $u^\varepsilon(t,x) \geq -(T+1)|H^*(0)| - M$ and $|u^\varepsilon|_0 \leq C$, for some constant C independent of ε.

The uniqueness of the viscosity solution is a direct consequence of the continuous dependence result as follows.

Lemma A.0.1 *For any $s \in (0, T+\varepsilon^2]$, let $u \in USC(\bar{Q}_s)$ be a bounded from above viscosity subsolution of (3.28) with coefficients σ^θ, b^θ and H^θ and $\bar{u} \in LSC(\bar{Q}_s)$ be a bounded from below viscosity supersolution of (3.28) with coefficients $\bar{\sigma}^\theta, \bar{b}^\theta$ and $\bar{H}^\theta$. Suppose that Assumption 3.2.1 holds for both sets of coefficients with respective constants M and $\bar{M}$, uniformly in $\theta \in \Theta^\varepsilon$, and that either $u(s,\cdot) \in \mathcal{C}_b^1(\mathbb{R}^d)$ or $\bar{u}(s,\cdot) \in \mathcal{C}_b^1(\mathbb{R}^d)$. Then, there exists a constant C, depending only on M, $\bar{M}$, $[u(s,\cdot)]_{\mathcal{C}^1}$ or $[\bar{u}(s,\cdot)]_{\mathcal{C}^1}$, and s, such that, for $(t,x) \in \bar{Q}_s$,*

$$u - \bar{u} \leq C\left(|(u-\bar{u})^+(s,\cdot)|_0 + \sup_{\theta\in\Theta^\varepsilon}\left\{|\sigma^\theta - \bar{\sigma}^\theta|_0 + |b^\theta - \bar{b}^\theta|_0\right\} + \sup_{\theta\in\Theta^\varepsilon}|H^\theta - \bar{H}^\theta|_0\right). \tag{A.2}$$

Proof. See section A.1. ∎

The x-regularity of u^ε follows easily from (A.2) by choosing $u = u^\varepsilon$, $\bar{u} = u^\varepsilon(\cdot, \cdot + e)$ and $s = T + \varepsilon^2$.

To get the t-regularity, we work as follows. Firstly, let $\rho(x)$ be a $\mathbb{R}_+$-valued smooth function with compact support $B(0,1)$ and mass 1 and introduce the sequence of mollifiers

$$\rho_\varepsilon(x) := \frac{1}{\varepsilon^d}\rho\left(\frac{x}{\varepsilon}\right). \tag{A.3}$$

For $0 \leq t < s \leq T+\varepsilon^2$, let $u_{\varepsilon'}$ be the unique bounded solution of (3.28) in Q_s with terminal condition $u_{\varepsilon'}(s,x) = u^\varepsilon(s,\cdot) * \rho_{\varepsilon'}(x)$, for some $\varepsilon' > 0$. It then follows from (A.2) that, for $(t,x) \in \bar{Q}_s$,

$$u^\varepsilon - u_{\varepsilon'} \leq C|(u^\varepsilon - u_{\varepsilon'})^+(s,\cdot)|_0 \leq C[u^\varepsilon(s,\cdot)]_{\mathcal{C}^1}\varepsilon' \leq C\varepsilon'.$$

Similarly, we also have $u_{\varepsilon'} - u^\varepsilon \leq C\varepsilon'$.

On the other hand, standard properties of mollifiers imply that $|D_x^j u_{\varepsilon'}(s,\cdot)|_0 \le C\varepsilon'^{1-|j|}$. Thus, for any $(\xi, x) \in Q_s$, we have

$$|\sup_{\theta\in\Theta^\varepsilon} g^\theta(\xi, x, D_x u_{\varepsilon'}(s,x), D_x^2 u_{\varepsilon'}(s,x))| \le C(\frac{1}{\varepsilon'}+1) =: C_{\varepsilon'}.$$

Define the functions $w^+_{\varepsilon'}(t,x) := u_{\varepsilon'}(s,x) + (s-t)C_{\varepsilon'}$ and $w^-_{\varepsilon'}(t,x) := u_{\varepsilon'}(s,x) - (s-t)C_{\varepsilon'}$. It then can be easily checked that $w^\pm_{\varepsilon'}$ are, respectively, bounded supersolution and subsolution of (3.28) in Q_s, with the same terminal condition $w^+_{\varepsilon'}(s,x) = w^-_{\varepsilon'}(s,x) = u_{\varepsilon'}(s,x)$. Thus, by (A.2), we have $w^-_{\varepsilon'}(t,x) \le u_{\varepsilon'}(t,x) \le w^+_{\varepsilon'}(t,x)$, for $(t,x) \in \bar{Q}_s$, which in turn implies that $|u_{\varepsilon'}(t,x) - u_{\varepsilon'}(s,x)| \le C_{\varepsilon'}|s-t|$. Therefore, choosing $\varepsilon' = \sqrt{|s-t|}$, we obtain that

$$\begin{aligned}
&|u^\varepsilon(t,x) - u^\varepsilon(s,x)| \\
&\le |u^\varepsilon(t,x) - u_{\varepsilon'}(t,x)| + |u_{\varepsilon'}(t,x) - u_{\varepsilon'}(s,x)| + |u_{\varepsilon'}(s,x) - u^\varepsilon(s,x)| \\
&\le 2C\varepsilon' + C_{\varepsilon'}|s-t| \le C(\varepsilon' + \frac{|s-t|}{\varepsilon'} + |s-t|) \le C\sqrt{|s-t|},
\end{aligned}$$

which, together with the boundedness and the x-regularity of u^ε, implies that $|u^\varepsilon|_1 \le C$.

Finally, note that $u(t,x)$ is also the bounded viscosity solution of (3.28) when $\sigma^\theta \equiv \sigma$, $b^\theta \equiv b$ and $H^\theta \equiv H$. Applying (A.2) once more and the regularity of σ, b, H and u^ε, we deduce that

$$\begin{aligned}
u^\varepsilon - u &\le C\left(|(u^\varepsilon - u)^+(T,\cdot)|_0 + \sup_{\theta\in\Theta^\varepsilon}\left\{|\sigma^\theta - \sigma|_0 + |b^\theta - b|_0\right\} + \sup_{\theta\in\Theta^\varepsilon}|H^\theta - H|_0\right) \\
&\le C\left(|u^\varepsilon(T,\cdot) - u^\varepsilon(T+\varepsilon^2,\cdot)|_0 + \varepsilon\right) \le C\varepsilon \quad \text{in } \bar{Q}_T.
\end{aligned}$$

Similarly, we also have $u - u^\varepsilon \le C\varepsilon$, and we easily conclude.

A.1 Proof of Lemma A.0.1

The proof follows a similar idea in the proof of Theorem A.1. in [39] with some modification and simplification. Fix $0 < s \le T$. By letting $v(t,x) = e^t u(t,x)$, $\bar{v}(t,x) = e^t \bar{u}(t,x)$ and $H'^\theta(t,x,p) = e^t H^\theta(t,x,e^{-t}p)$, $\bar{H}'^\theta(t,x,p) = e^t \bar{H}^\theta(t,x,e^{-t}p)$, it is simple to derive that $v \in USC([0,s]\times\mathbb{R}^d)$ and $\bar{v} \in LSC([0,s]\times\mathbb{R}^d)$ are bounded above viscosity subsolution and bounded below supersolution of the following PDE with coefficients $\{\sigma^\theta, b^\theta, H^\theta\}$ and $\{\bar{\sigma}^\theta, \bar{b}^\theta, \bar{H}^\theta\}$ respectively:

$$-\partial_t v + v + \sup_{\theta\in\Theta^\varepsilon} G'^\theta(t,x,D_x v, D_x^2 v) = 0.$$

with

$$G'^\theta(t,x,p,X) = -\frac{1}{2}\mathrm{tr}\left(\sigma^\theta\sigma^{\theta^T}(t,x)X\right) - b^\theta(t,x)\cdot p + H'^\theta(t,x,p).$$

We now use a doubling variables argument to derive an upper bound for $v - \bar{v}$ and then will derive for $u - \bar{u}$ by using back-substitution.

To continue, we define in $[0,s]\times\mathbb{R}^d\times\mathbb{R}^d$ that $\phi(t,x,y) := e^{\lambda(s-t)}\alpha|x-y|^2 + \epsilon(|x|^2+|y|^2)$ and $\psi(t,x,y) := v(t,x) - \bar{v}(t,y) - \phi(t,x,y)$, where $\lambda, \alpha, \epsilon > 0$

are positive constants. Then we let $m^s_{\alpha,\epsilon} = \sup_{\mathbb{R}^d\times\mathbb{R}^d} \psi(s,x,y)^+$ and $m_{\lambda,\alpha,\epsilon} = \sup_{[0,s]\times\mathbb{R}^d\times\mathbb{R}^d} \psi(t,x,y) - m^s_{\alpha,\epsilon}$. Since $v, -\bar{v}$ are bounded from above and $\phi \geq 0$, there exists $t_0 \in [0,s]$ and $x_0, y_0 \in \mathbb{R}^d$ depending on λ, α and ϵ such that $\psi(t_0,x_0,y_0) = \sup_{[0,s]\times\mathbb{R}^d\times\mathbb{R}^d} \psi(t,x,y)$. Note that by letting $y = x$, we have

$$m^s_{\alpha,\epsilon} + m_{\lambda,\alpha,\epsilon} = \psi(t_0,x_0,y_0) \geq v(t,x) - \bar{v}(t,x) - 2\epsilon|x|^2 \tag{A.4}$$

for any $(t,x) \in [0,s]\times\mathbb{R}^d$. We now try to derive the upper bound for $m^s_{\alpha,\epsilon}$ and $m_{\lambda,\alpha,\epsilon}$.

Since $u(s,\cdot)$ or $\bar{u}(s,\cdot)$ is in $\mathcal{C}^1_b(\mathbb{R}^d)$, without loss of generality we assume $[\bar{u}(s,\cdot)]_{\mathcal{C}^1} \leq C$ for some constant C, then for any $x, y \in \mathbb{R}^d$,

$$\begin{aligned}
\psi(s,x,y) &\leq v(s,x) - \bar{v}(s,y) - \alpha|x-y|^2 \\
&\leq |(v-\bar{v})^+(s,\cdot)|_0 + [\bar{v}(s,\cdot)]_{\mathcal{C}^1}|x-y| - \alpha|x-y|^2 \\
&= |(v-\bar{v})^+(s,\cdot)|_0 + e^s[\bar{u}(s,\cdot)]_{\mathcal{C}^1}|x-y| - \alpha|x-y|^2 \\
&\leq |(v-\bar{v})^+(s,\cdot)|_0 + C\alpha^{-1}
\end{aligned}$$

where the constant C depends only on $[\bar{u}(s,\cdot)]_{\mathcal{C}^1}$ and s, and the last inequality follows from that $\sup_{r\geq 0}(ar - br^2) = a^2/4b$ for any $a, b > 0$. Thus, we get the upper bound for $m^s_{\alpha,\epsilon}$:

$$m^s_{\alpha,\epsilon} \leq |(v-\bar{v})^+(s,\cdot)|_0 + C\alpha^{-1}. \tag{A.5}$$

On the other hand, we assume that $m_{\lambda,\alpha,\epsilon} > 0$ and derive its (positive) upper bound. Of course this upper bound still holds for $m_{\lambda,\alpha,\epsilon} \leq 0$. Follow this assumption, we have $t_0 < s$, since otherwise, $m_{\lambda,\alpha,\epsilon} = \sup_{\mathbb{R}^d\times\mathbb{R}^d} \psi(s,x,y) - m^s_{\alpha,\epsilon} \leq 0$. Then we can apply the parabolic version of Crandall-Ishii's Lemma (see Lemma 2.2.12) to get that there exist $a, b \in \mathbb{R}$ and $X, Y \in \mathcal{S}_d$ such that $(a, D_x\phi(t_0,x_0,y_0), X) \in \bar{\mathcal{P}}^{2,+}v(t_0,x_0)$ and $(b, -D_y\phi(t_0,x_0,y_0), Y) \in \bar{\mathcal{P}}^{2,-}\bar{v}(t_0,y_0)$ with $a - b = \phi_t(t_0,x_0,y_0)$ and the following inequality holds

$$\begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq 3e^{\lambda(s-t_0)}\alpha \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + 3\epsilon \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}. \tag{A.6}$$

Then by the definitions of viscosity sub- and supersolutions, we have

$$-a + v(t_0,x_0) + \sup_{\theta\in\Theta^\varepsilon} G'^\theta(t_0,x_0,D_x\phi(t_0,x_0,y_0),X) \leq 0$$

and

$$-b + \bar{v}(t_0,y_0) + \sup_{\theta\in\Theta^\varepsilon} \bar{G}'^\theta(t_0,y_0,-D_y\phi(t_0,x_0,y_0),Y) \geq 0$$

Using the fact that $\sup(A) - \sup(B) \leq \sup(A - B)$, we then have

$$\begin{aligned}
0 \leq & -\lambda e^{\lambda(s-t_0)}\alpha|x_0 - y_0|^2 - v(t_0, x_0) + \bar{v}(t_0, y_0) \\
& + \sup_{\theta \in \Theta^\varepsilon} \left\{ \frac{1}{2}\mathrm{tr}[\sigma^\theta \sigma^{\theta^T}(t_0, x_0)X - \bar{\sigma}^\theta \bar{\sigma}^{\theta^T}(t_0, y_0)Y] \right. \\
& - \bar{b}^\theta(t_0, y_0) \cdot (2e^{\lambda(s-t_0)}\alpha(x_0 - y_0) - 2\epsilon y_0) \\
& + b^\theta(t_0, x_0) \cdot (2e^{\lambda(s-t_0)}\alpha(x_0 - y_0) + 2\epsilon x_0) \\
& + \bar{H}'^\theta(t_0, y_0, 2e^{\lambda(s-t_0)}\alpha(x_0 - y_0) - 2\epsilon y_0) \\
& - \left. H'^\theta(t_0, x_0, 2e^{\lambda(s-t_0)}\alpha(x_0 - y_0) + 2\epsilon x_0) \right\} \qquad \text{(A.7)}
\end{aligned}$$

Now we use the Assumption (3.2.1) for all the coefficients. By the inequality (A.6) and the fact that $(s+t)^2 \leq 2(s^2 + t^2)$ for $s, t \in \mathbb{R}$, we get

$$\begin{aligned}
& \mathrm{tr}[\sigma^\theta \sigma^{\theta^T}(t_0, x_0)X - \bar{\sigma}^\theta \bar{\sigma}^{\theta^T}(t_0, y_0)Y] \\
\leq & \, 6e^{\lambda(s-t_0)}\alpha \left\{ |\sigma^\theta(t_0, x_0) - \bar{\sigma}^\theta(t_0, x_0)|^2 + |\bar{\sigma}^\theta(t_0, x_0) - \bar{\sigma}^\theta(t_0, y_0)|^2 \right\} \\
& + 3\epsilon \left\{ |\sigma^\theta(t_0, x_0)|^2 + |\bar{\sigma}^\theta(t_0, y_0)|^2 \right\} \\
\leq & \, 6e^{\lambda(s-t_0)}\alpha \left\{ |\sigma^\theta - \bar{\sigma}^\theta|_0^2 + [\bar{\sigma}^\theta]^2_{\mathcal{C}^{1/2,1}} |x_0 - y_0|^2 \right\} \\
& + 3\epsilon \left\{ |\sigma^\theta|_0^2 + |\bar{\sigma}^\theta|_0^2 \right\}.
\end{aligned}$$

Furthermore, we have the following estimates

$$\begin{aligned}
& (b^\theta(t_0, x_0) - \bar{b}^\theta(t_0, y_0)) \cdot (x_0 - y_0) \\
\leq & \, \frac{1}{2}\left(|b^\theta(t_0, x_0) - \bar{b}^\theta(t_0, x_0)|^2 + |x_0 - y_0|^2 \right) \\
& + |\bar{b}^\theta(t_0, x_0) - \bar{b}^\theta(t_0, y_0)||x_0 - y_0| \\
\leq & \, \frac{1}{2}\left(|b^\theta - \bar{b}^\theta|_0^2 + |x_0 - y_0|^2 \right) + [\bar{b}^\theta]_{\mathcal{C}^{1/2,1}} |x_0 - y_0|^2,
\end{aligned}$$

$$\begin{aligned}
b^\theta(t_0, x_0) \cdot x_0 \leq |b^\theta(t_0, x_0)||x_0| &\leq ([b^\theta]_{\mathcal{C}^{1/2,1}}|x_0| + |b^\theta(t_0, 0)|)|x_0| \\
&\leq M(1 + |x_0|)^2 \leq 2M(1 + |x_0|^2),
\end{aligned}$$

similarly

$$\bar{b}^\theta(t_0, y_0) \cdot y_0 \leq 2\bar{M}(1 + |y_0|^2),$$

$$\bar{v}(t_0, y_0) - v(t_0, x_0) \leq -\psi(t_0, x_0, y_0) \leq -m_{\lambda,\alpha,\epsilon},$$

and

$$\begin{aligned}
& \bar{H}'^\theta(t_0, y_0, 2e^{\lambda(s-t_0)}\alpha(x_0 - y_0) - 2\epsilon y_0) - H'^\theta(t_0, x_0, 2e^{\lambda(s-t_0)}\alpha(x_0 - y_0) - 2\epsilon x_0) \\
\leq & \, |H'^\theta - \bar{H}'^\theta|_0 + C|x_0 - y_0| \\
& + H'^\theta(t_0, x_0, 2e^{\lambda(s-t_0)}\alpha(x_0 - y_0) - 2\epsilon y_0) - H'^\theta(t_0, x_0, 2e^{\lambda(s-t_0)}\alpha(x_0 - y_0) - 2\epsilon x_0).
\end{aligned}$$

Plugging in all these estimates into inequality (A.7) yields

$$
\begin{aligned}
m_{\lambda,\alpha,\epsilon} \leq\ & 3e^{\lambda(s-t_0)}\alpha \sup_{\theta\in\Theta^\varepsilon}\left\{|\sigma^\theta-\bar{\sigma}^\theta|_0^2+|b^\theta-\bar{b}^\theta|_0^2\right\}+(k-\lambda)e^{\lambda(s-t_0)}\alpha|x_0-y_0|^2 \\
&+ C|x_0-y_0| + \sup_{\theta\in\Theta^\varepsilon}\Big\{|H'^\theta-\bar{H}'^\theta|_0 + H'^\theta(t_0,x_0,2e^{\lambda(s-t_0)}\alpha(x_0-y_0)-2\epsilon y_0) \\
&- H'^\theta(t_0,x_0,2e^{\lambda(s-t_0)}\alpha(x_0-y_0)+2\epsilon x_0)\Big\} \\
&+ C(1+|x_0|^2+|y_0|^2)\epsilon \qquad \text{(A.8)}
\end{aligned}
$$

where $k = 3\bar{M}^2 + 2\bar{M} + 1$ is some positive constant. Then we combine the upper bounds (A.5) and (A.8) for $m^s_{\alpha,\epsilon}$ and $m_{\lambda,\alpha,\epsilon}$ and plug them into (A.4) to get

$$
\begin{aligned}
v(t,x)-\bar{v}(t,x) \leq\ & |(v-\bar{v})^+(s,\cdot)|_0 + 3e^{\lambda(s-t_0)}\alpha \sup_{\theta\in\Theta^\varepsilon}\left\{|\sigma^\theta-\bar{\sigma}^\theta|_0^2+|b^\theta-\bar{b}^\theta|_0^2\right\} \\
&+ (k-\lambda)e^{\lambda(s-t_0)}\alpha|x_0-y_0|^2 + C|x_0-y_0| + C\alpha^{-1} \\
&+ \sup_{\theta\in\Theta^\varepsilon}\Big\{|H'^\theta-\bar{H}'^\theta|_0 + H'^\theta(t_0,x_0,2e^{\lambda(s-t_0)}\alpha(x_0-y_0)-2\epsilon y_0) \\
&- H'^\theta(t_0,x_0,2e^{\lambda(s-t_0)}\alpha(x_0-y_0)+2\epsilon x_0)\Big\} \\
&+ C(1+|x_0|^2+|y_0|^2)\epsilon + 2\epsilon|x|^2
\end{aligned}
$$

Note that this estimate holds for any $\lambda, \alpha, \epsilon > 0$. We then try to select appropriate value for them (or take limit) to draw our conclusion. Firstly we may choose $\lambda = k+1$ and follow again that $\sup_{r\geq 0}(ar - br^2) = a^2/4b$ to get rid of the $|x_0 - y_0|$ term. Then, by standard arguments, we know that for any fixed λ and α, $\lim_{\epsilon\to 0}\epsilon(|x_0|^2+|y_0|^2) = 0$, which also implies $\lim_{\epsilon\to 0}\epsilon(x_0+y_0) = 0$, together with the continuity of H'^θ, by letting $\epsilon \to 0$, we further get

$$
\begin{aligned}
v(t,x)-\bar{v}(t,x) \leq\ & |(v-\bar{v})^+(s,\cdot)|_0 + 3e^{(k+1)(s-t_0)}\alpha \sup_{\theta\in\Theta^\varepsilon}\left\{|\sigma^\theta-\bar{\sigma}^\theta|_0^2+|b^\theta-\bar{b}^\theta|_0^2\right\} \\
&+ (C+Ce^{-(k+1)(s-t_0)})\alpha^{-1} + \sup_{\theta\in\Theta^\varepsilon}|H'^\theta-\bar{H}'^\theta|_0.
\end{aligned}
$$

Note that $\min_{\alpha>0}(C_1\alpha + C_2\alpha^{-1}) = 2(C_1C_2)^{1/2}$ for any $C_1, C_2 > 0$, we can choose the α minimising the right hand side to get

$$
\begin{aligned}
v(t,x)-\bar{v}(t,x) \leq\ & |(v-\bar{v})^+(s,\cdot)|_0 + C\sup_{\theta\in\Theta^\varepsilon}\left\{|\sigma^\theta-\bar{\sigma}^\theta|_0+|b^\theta-\bar{b}^\theta|_0\right\} \\
&+ \sup_{\theta\in\Theta^\varepsilon}|H'^\theta-\bar{H}'^\theta|_0.
\end{aligned}
$$

where we used that $(s^2+t^2)^{1/2} \leq |s|+|t|$ for any $s, t \in \mathbb{R}$. Finally, the conclusion follows by back-substituting v, $\bar{v}$, H'^θ, and $\bar{H}'^\theta$ by u and $\bar{u}$, H^θ, and $\bar{H}^\theta$ respectively.

Appendix B

Proofs of Proposition 4.1.2 and 4.2.3

The proofs follow from essentially the same arguments in the proofs of Proposition 3.2.2 and 3.3.3 in Appendix A, with some modifications to incorporate the obstacle term f. Since equation (4.1) is a special case (choosing $\varepsilon = 0$) of the equation (4.10), we omit the proof of Proposition 4.1.2 and only prove Proposition 4.2.3.

We first show that there exists a bounded solution to (4.10). Similar to (A.1), we rewrite (4.10) as

$$\begin{cases} \max\{-\partial_t u^\varepsilon + \sup_{\theta\in\Theta^\varepsilon, q\in\mathbb{R}^d} \mathcal{L}^{\theta,q}\left(t,x,D_x u^\varepsilon, D_x^2 u^\varepsilon\right), u^\varepsilon - f(t,x)\} = 0 & \text{in } Q_{T+\varepsilon^2}; \\ u^\varepsilon(T+\varepsilon^2, x) = U(x) & \text{in } \mathbb{R}^d, \end{cases} \tag{B.1}$$

whose bounded viscosity solution can be represented by the value function of a mixed stochastic control and optimal stopping problem on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$:

$$\begin{aligned} u^\varepsilon(t,x) = & \inf_{\theta\in\Theta^\varepsilon[t,T+\varepsilon^2], q\in\mathbb{H}^2[t,T+\varepsilon^2]} \inf_{\nu\in\mathcal{T}_{[t,T+\varepsilon^2]}} \mathbb{E}\left[\int_t^\nu L^{\theta_s}\left(s, X_s^{t,x;\theta,q}, q_s\right) ds\right. \\ & + \left. f(\nu, X_\nu^{t,x;\theta,q}) \mathbf{1}_{\{\nu<T+\varepsilon^2\}} + U\left(X_{T+\varepsilon^2}^{t,x;\theta,q}\right) \mathbf{1}_{\{\nu=T+\varepsilon^2\}} | \mathcal{F}_t\right], \end{aligned}$$

with the controlled state equation

$$dX_s^{t,x;\theta,q} = \left(b^{\theta_s}(s, X_s^{t,x;\theta,q}) - q_s\right) ds + \sigma^{\theta_s}\left(s, X_s^{t,x;\theta,q}\right) dW_s,$$

where $\Theta^\varepsilon[t, T+\varepsilon^2]$ is the space of Θ^ε-valued progressively measurable processes (τ_s, e_s), $\mathbb{H}^2[t, T+\varepsilon^2]$ is the space of square-integrable progressively measurable processes q_s, for $s \in [t, T+\varepsilon^2]$, $\mathcal{T}_{[t,T+\varepsilon^2]}$ is the collection of all $\mathbb{F}$-stopping times with values in $[t, T+\varepsilon^2]$, and W is an d-dimensional Brownian motion with its augmented filtration $\mathbb{F}$.

To find an upper bound of u^ε, we choose an arbitrary perturbation parameter process $\theta \in \Theta^\varepsilon[t, T+\varepsilon]$, an arbitrary stopping time $\nu \in \mathcal{T}_{[t,T+\varepsilon^2]}$, and a particular control $\hat{q}$ with $\hat{q}_s \equiv 0$. Then, Proposition 3.2.3 (ii) yields

$$\begin{aligned}u^\varepsilon(t,x) &\le \mathbb{E}\left[\int_t^\nu L^{\theta_s}(s, X_s^{t,x;\theta,\hat{q}}, 0)ds + |f|_0 \mathbf{1}_{\{\nu<T+\varepsilon^2\}} + |U|_0 \mathbf{1}_{\{\nu=T+\varepsilon^2\}} |\mathcal{F}_t\right] \\ &\le (T+\varepsilon^2-t)|L^*(0)| + M \le C.\end{aligned}$$

For the lower bound, we use again Proposition 3.2.3 (ii) to obtain that $L_*(q) \ge -H^*(0) \ge -|H^*(0)|$, for any $q \in \mathbb{R}^d$. In turn, for any $(\theta, q, \nu) \in \Theta^\varepsilon[t, T+\varepsilon^2] \times \mathbb{H}^2[t, T+\varepsilon^2] \times \mathcal{T}_{[t,T+\varepsilon^2]}$,

$$\begin{aligned}&\mathbb{E}\left[\int_t^\nu L^{\theta_s}\left(s, X_s^{t,x;\theta,q}, q_s\right) ds + f(\nu, X_\nu^{t,x;\theta,q})\mathbf{1}_{\{\nu<T+\varepsilon^2\}}\right. \\ &+ \left. U\left(X_{T+\varepsilon^2}^{t,x;\theta,q}\right)\mathbf{1}_{\{\nu=T+\varepsilon^2\}}|\mathcal{F}_t\right] \\ &\ge \mathbb{E}\left[\int_t^\nu L_*(q_s)ds|\mathcal{F}_t\right] - M \ge -(T+\varepsilon^2-t)|H^*(0)| - M \ge -C,\end{aligned}$$

and, thus, $u^\varepsilon(t,x) \ge -C$ and $|u^\varepsilon|_0 \le C$, for some constant C independent of ε.

The uniqueness of the viscosity solution is a direct consequence of the following continuous dependence result, whose proof follows along similar arguments as in Theorem A.0.1, and is thus omitted.

Lemma B.0.1 *For any $s \in (0, T+\varepsilon^2]$, let $u \in USC(\bar{Q}_s)$ be a bounded from above viscosity subsolution of (4.10) with coefficients $\{\sigma^\theta, b^\theta\, H^\theta, f\}$, and $\bar{u} \in LSC(\bar{Q}_s)$ be a bounded from below viscosity supersolution of (4.10) with coefficients $\{\bar{\sigma}^\theta, \bar{b}^\theta, \bar{H}^\theta, \bar{f}\}$. Suppose that Assumption 3.2.1 and 4.1.1 hold for both sets of coefficients with respective constants M and $\bar{M}$, uniformly in $\theta \in \Theta^\varepsilon$, and that either $u(s,\cdot) \in \mathcal{C}_b^1(\mathbb{R}^d)$ or $\bar{u}(s,\cdot) \in \mathcal{C}_b^1(\mathbb{R}^d)$. Then, there exists a constant C, depending only on M, $\bar{M}$, $[u(s,\cdot)]_1$ or $[\bar{u}(s,\cdot)]_1$, and s, such that in $\bar{Q}_s$,*

$$\begin{aligned}u - \bar{u} \le C\Bigg(&|(u-\bar{u})^+(s,\cdot)|_0 + \sup_{\theta\in\Theta^\varepsilon}\left\{|\sigma^\theta - \bar{\sigma}^\theta|_0 + |b^\theta - \bar{b}^\theta|_0\right\} \\ &+ \sup_{\theta\in\Theta^\varepsilon}|H^\theta - \bar{H}^\theta|_0 + |f-\bar{f}|_0\Bigg). \qquad \text{(B.2)}\end{aligned}$$

We now continue the proof of Proposition 4.2.3. The x-regularity of u^ε then follows easily from (B.2) by choosing $u = u^\varepsilon$, $\bar{u} = u^\varepsilon(\cdot, \cdot + e)$ and $s = T+\varepsilon^2$. To get the t-regularity, let $\rho(x)$ be a $\mathbb{R}_+$-valued smooth function with compact support $B(0,1)$ and mass 1, and for $\varepsilon > 0$, let $\rho_\varepsilon(x) := \frac{1}{\varepsilon^d}\rho\left(\frac{x}{\varepsilon}\right)$ be a sequence of mollifiers. Next, fix any (t,s) such that $0 \le t < s \le T+\varepsilon^2$ and let $u_{\varepsilon'}$ be the unique bounded solution of (4.10) in Q_s with terminal condition $u_{\varepsilon'}(s,x) = u^\varepsilon(s,\cdot) * \rho_{\varepsilon'}(x)$, for some $\varepsilon' > 0$ that shall be decided later. It then follows from (B.2) that, in $\bar{Q}_s$,

$$u^\varepsilon - u_{\varepsilon'} \le C|(u^\varepsilon - u_{\varepsilon'})^+(s,\cdot)|_0 \le C[u^\varepsilon(s,\cdot)]_{\mathcal{C}^1}\varepsilon' \le C\varepsilon'.$$

Similarly, we also have $u_{\varepsilon'} - u^\varepsilon \le C\varepsilon'$.

On the other hand, standard properties of mollifiers imply that $|D_x^j u_{\varepsilon'}(s,\cdot)|_0 \le$

$C\varepsilon'^{1-|j|}$. Thus, for any $(\xi, x) \in Q_s$, we have

$$|\sup_{\theta \in \Theta^\varepsilon} g^\theta(\xi, x, D_x u_{\varepsilon'}(s,x), D_x^2 u_{\varepsilon'}(s,x))| \leq C(\frac{1}{\varepsilon'} + 1) =: C_{\varepsilon'}.$$

Define two functions $w_{\varepsilon'}^+(t,x) := u_{\varepsilon'}(s,x) + (s-t)C_{\varepsilon'}$ and $w_{\varepsilon'}^-(t,x) := u_{\varepsilon'}(s,x) - (s-t)C_{\varepsilon'}$. It then can be easily checked that the function $w_{\varepsilon'}^+(t,x)$ is a bounded supersolution of (4.10) in Q_s. Thus, by (B.2), we have in $\bar{Q}_s$, $u_{\varepsilon'} - w_{\varepsilon'}^+ \leq 0$, which impies that

$$u_{\varepsilon'}(t,x) - u_{\varepsilon'}(s,x) \leq C_{\varepsilon'}(s-t).$$

We then construct a bounded subsolution of (4.10) in Q_s based on $w_{\varepsilon'}^-$. Since $u_{\varepsilon'}(s,x) \leq u^\varepsilon(s,x) + C\varepsilon' \leq f(s,x) + C\varepsilon'$, we obtain that for any $(\xi, x) \in Q_s$,

$$\begin{aligned} w_{\varepsilon'}^-(\xi,x) - f(\xi,x) &= u_{\varepsilon'}(s,x) - C_{\varepsilon'}(s-\xi) - f(\xi,x) \\ &\leq f(s,x) - f(\xi,x) - C_{\varepsilon'}(s-\xi) + C\varepsilon' \\ &\leq M\sqrt{s-\xi} - C_{\varepsilon'}(s-\xi) + C\varepsilon' \\ &\leq \frac{M^2}{4C_{\varepsilon'}} + C\varepsilon' \leq C\varepsilon', \end{aligned} \tag{B.3}$$

where we used $\sup_{r\geq 0}(ar - br^2) = a^2/4b$ for any $a, b > 0$, and $\frac{1}{C_{\varepsilon'}} \leq C\varepsilon'$. This implies that $w_{\varepsilon'}^- - C\varepsilon'$ is a bounded subsolution of (4.10) in Q_s. By (B.2), we then have in $\bar{Q}_s$, $w_{\varepsilon'}^- - C\varepsilon' - u_{\varepsilon'} \leq 0$, which impies that

$$u_{\varepsilon'}(s,x) - u_{\varepsilon'}(t,x) \leq C_{\varepsilon'}(s-t) + C\varepsilon'.$$

Therefore, choosing $\varepsilon' = \sqrt{|s-t|}$, we obtain that

$$\begin{aligned} &|u^\varepsilon(t,x) - u^\varepsilon(s,x)| \\ &\leq |u^\varepsilon(t,x) - u_{\varepsilon'}(t,x)| + |u_{\varepsilon'}(t,x) - u_{\varepsilon'}(s,x)| + |u_{\varepsilon'}(s,x) - u^\varepsilon(s,x)| \\ &\leq 2C\varepsilon' + C_{\varepsilon'}|s-t| \leq C(\varepsilon' + \frac{|s-t|}{\varepsilon'} + |s-t|) \leq C\sqrt{|s-t|}, \end{aligned}$$

which, together with the boundedness and the x-regularity of u^ε, implies that $|u^\varepsilon|_1 \leq C$.

Finally, note that $u(t,x)$ is also the bounded viscosity solution of (4.10) when $\sigma^\theta \equiv \sigma$, $b^\theta \equiv b$ and $H^\theta \equiv H$. Applying (B.2) once more and the regularity of σ, b, H and u^ε, we deduce that in $\bar{Q}_T$,

$$\begin{aligned} u^\varepsilon - u &\leq C\left(|(u^\varepsilon - u)^+(T,\cdot)|_0 + \sup_{\theta\in\Theta^\varepsilon}\left\{|\sigma^\theta - \sigma|_0 + |b^\theta - b|_0\right\} + \sup_{\theta\in\Theta^\varepsilon}|H^\theta - H|_0\right) \\ &\leq C\left(|u^\varepsilon(T,\cdot) - u^\varepsilon(T+\varepsilon^2,\cdot)|_0 + \varepsilon\right) \leq C\varepsilon, \end{aligned}$$

Similarly, we also have $u - u^\varepsilon \leq C\varepsilon$, and we conclude.

Appendix C

Proof of Proposition 6.4.2

The convergence rate $\frac{\alpha}{2}$ follows immediately from Theorem 5.1.4(iii). To establish the other convergence rate $\frac{1}{4}$, we only need to prove the following consistency error estimate

$$\begin{aligned}\mathcal{E}(\Delta,\psi) &:= \left|\frac{\mathbf{S}(\Delta)\psi-\psi}{\Delta}-G(D\psi,D^2\psi)\right|_0\\ &\leq C\Delta(|D^4\psi|_0+|D^3\psi|_0+|D^2\psi|_0),\end{aligned}\tag{C.1}$$

for any test function $\psi\in\mathcal{C}_b^\infty(\mathbb{R})$. Note that (C.1) is a refinement of the consistency error estimate in Proposition 5.3.1. The rest of the proof then follows along a similar argument and procedure used in the proof of Theorem 5.1.4. To establish (C.1), with $Y=r-\frac{1}{2}X^2$, we have

$$\begin{aligned}&\mathbf{S}(\Delta)\psi(x)-\psi(x)-\Delta G(D\psi(x),D^2\psi(x))\\ &\leq\hat{\mathbb{E}}[\psi(x+\sqrt{\Delta}X+\Delta Y)-\psi(x)-\Delta D\psi(x)Y-\frac{1}{2}\Delta D^2\psi(x)X^2]\\ &\leq\hat{\mathbb{E}}[\psi(x+\sqrt{\Delta}X)-\psi(x)-\frac{1}{2}D^2\psi(x)\Delta X^2]\\ &\quad+\hat{\mathbb{E}}[\psi(x+\sqrt{\Delta}X+\Delta Y)-\psi(x+\sqrt{\Delta}X)-D\psi(x+\sqrt{\Delta}X)\Delta Y]\\ &\quad+\hat{\mathbb{E}}[D\psi(x+\sqrt{\Delta}X)\Delta Y-D\psi(x)\Delta Y]:=(I)+(II)+(III).\end{aligned}$$

Since $\hat{\mathbb{E}}[X]=\hat{\mathbb{E}}[-X]=\hat{\mathbb{E}}[X^3]=\hat{\mathbb{E}}[-X^3]=0$, we have

$$\begin{aligned}(I)&=\hat{\mathbb{E}}\left[\sqrt{\Delta}D\psi(x)X+\int_x^{x+\sqrt{\Delta}X}\int_x^s(D^2\psi(u)-D^2\psi(x))duds\right]\\ &=\hat{\mathbb{E}}\left[\int_x^{x+\sqrt{\Delta}X}\int_x^s\int_x^u D^3\psi(p)dpduds\right]\\ &=\hat{\mathbb{E}}\left[\Delta^{3/2}D^3\psi(x)X^3+\int_x^{x+\sqrt{\Delta}X}\int_x^s\int_x^u(D^3\psi(p)-D^3\psi(x))dpduds\right]\\ &=\hat{\mathbb{E}}\left[\int_x^{x+\sqrt{\Delta}X}\int_x^s\int_x^u(D^3\psi(p)-D^3\psi(x))dpduds\right]\\ &\leq|D^4\psi|_0\hat{\mathbb{E}}\left|\int_x^{x+\sqrt{\Delta}X}\int_x^s\int_x^u|p-x|dpduds\right|\leq\Delta^2|D^4\psi|_0M_X^4.\end{aligned}$$

Likewise, Taylor's expansion yields that

$$
\begin{aligned}
(II) &= \hat{\mathbb{E}}\left[\int_0^1 (1-s)D^2\psi(x+\sqrt{\Delta}X+s\Delta Y)\Delta^2 Y^2 ds\right] \\
&\le \hat{\mathbb{E}}\left[\int_0^1 (1-s)ds|D^2\psi|_0\Delta^2 Y^2\right] \le \frac{1}{2}\Delta^2|D^2\psi|_0 M_Y^2,
\end{aligned}
$$

and

$$
\begin{aligned}
(III) &= \hat{\mathbb{E}}\left[\int_0^1 D^2\psi(x+s\sqrt{\Delta}X)\sqrt{\Delta}X ds\Delta Y\right] \\
&= \hat{\mathbb{E}}\left[\int_0^1 [D^2\psi(x+s\sqrt{\Delta}X)-D^2\psi(x)]ds\Delta^{\frac{3}{2}}XY + D^2\psi(x)\Delta^{\frac{3}{2}}XY\right] \\
&\le \hat{\mathbb{E}}\left[\int_0^1 [D^2\psi(x+s\sqrt{\Delta}X)-D^2\psi(x)]ds\Delta^{\frac{3}{2}}XY\right] + \hat{\mathbb{E}}\left[D^2\psi(x)\Delta^{\frac{3}{2}}XY\right] \\
&= \hat{\mathbb{E}}\left[\int_0^1\int_0^1 D^3\psi(x+us\sqrt{\Delta}X)s\sqrt{\Delta}X duds\Delta^{\frac{3}{2}}XY\right] \\
&\le \hat{\mathbb{E}}\left[\int_0^1\int_0^1 sduds|D^3\psi|_0\Delta^2|X|^2|Y|\right] \le \frac{1}{4}\Delta^2|D^3\psi|_0(M_X^4+M_Y^2),
\end{aligned}
$$

where we also used the fact that $\hat{\mathbb{E}}[XY]=\hat{\mathbb{E}}[-XY]=0$. The consistency error estimate (C.1) then follows by combining (I)-(III).

Bibliography

[1] M. Avellaneda, A. Levy, and A. Parás. Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance*, 2(2):73–88, 1995.

[2] G. Barles and E. R. Jakobsen. On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 36(1):33–54, 2002.

[3] G. Barles and E. R. Jakobsen. Error bounds for monotone approximation schemes for Hamilton–Jacobi–Bellman equations. *SIAM journal on numerical analysis*, 43(2):540–558, 2005.

[4] G. Barles and E. R. Jakobsen. Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations. *Mathematics of Computation*, 76(260):1861–1893, 2007.

[5] G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic analysis*, 4(3):271–283, 1991.

[6] E. Bayraktar and A. Fahim. A stochastic approximation for fully nonlinear free boundary parabolic problems. *Numerical Methods for Partial Differential Equations*, 30(3):902–929, 2014.

[7] A. C. Berry. The accuracy of the Gaussian approximation to the sum of independent variates. *Transactions of the american mathematical society*, 49(1):122–136, 1941.

[8] O. Bokanowski, S. Maroso, and H. Zidani. Some convergence results for Howard's algorithm. *SIAM Journal on Numerical Analysis*, 47(4):3001–3026, 2009.

[9] O. Bokanowski, A. Picarelli, and H. Zidani. Dynamic programming and error estimates for stochastic control problems with maximum cost. *Applied Mathematics & Optimization*, 71(1):125–163, 2015.

[10] B. Bouchard and N. Touzi. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Processes and their applications*, 111(2):175–206, 2004.

[11] B. Bouchard and N. Touzi. Weak dynamic programming principle for viscosity solutions. *SIAM Journal on Control and Optimization*, 49(3):948–962, 2011.

[12] L. A. Caffarelli and P. E. Souganidis. A rate of convergence for monotone finite difference approximations to fully nonlinear, uniformly elliptic pdes. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 61(1):1–17, 2008.

[13] F. Camilli and M. Falcone. An approximation scheme for the optimal control of diffusion processes. *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique*, 29(1):97–122, 1995.

[14] J. F. Chassagneux, A. Richou, et al. Numerical simulation of quadratic BSDEs. *The Annals of Applied Probability*, 26(1):262–304, 2016.

[15] W. F. Chong, Y. Hu, G. Liang, and T. Zariphopoulou. An ergodic BSDE approach to forward entropic risk measures: representation and large-maturity behavior. *Finance and Stochastics*, 23(1):239–273, 2019.

[16] S. Cohen, S. Ji, and S. Peng. Sublinear expectations and martingales in discrete time. *arXiv preprint arXiv:1104.5390*, 2011.

[17] M. G. Crandall, H. Ishii, and P.-L. Lions. Uniqueness of viscosity solutions of hamilton-jacobi equations revisited. *Journal of the Mathematical Society of Japan*, 39(4):581–596, 1987.

[18] M. G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bulletin of the American mathematical society*, 27(1):1–67, 1992.

[19] M. G. Crandall and P.-L. Lions. Viscosity solutions of hamilton-jacobi equations. *Transactions of the American mathematical society*, 277(1):1–42, 1983.

[20] M. G. Crandall and P.-L. Lions. On existence and uniqueness of solutions of Hamilton-Jacobi equations. *Nonlinear Analysis: Theory, Methods & Applications*, 10(4):353–370, 1986.

[21] K. Debrabant and E. Jakobsen. Semi-lagrangian schemes for linear and fully non-linear diffusion equations. *Mathematics of Computation*, 82(283):1433–1462, 2013.

[22] F. Delbaen, Y. Hu, and X. Bao. Backward SDEs with superquadratic growth. *Probability Theory and Related Fields*, 150(1-2):145–192, 2011.

[23] R. Dumitrescu, C. Reisinger, and Y. Zhang. Approximation schemes for mixed optimal stopping and control problems with nonlinear expectations and jumps. *Applied Mathematics & Optimization*, pages 1–43, 2019.

[24] C. G. Esseen. On the Liapunoff limit of error in the theory of probability. *Arkiv for matematik, astronomi och fysik, A: 1–19*, 1942.

[25] L. C. Evans. A second order elliptic equation with gradient constraint. *Communications in Partial Differential Equations*, 4(5):555–572, 1979.

[26] L. C. Evans. Partial differential equations and monge-kantorovich mass transfer. *Current developments in mathematics*, 1997(1):65–126, 1997.

[27] L. C. Evans and A. Friedman. Optimal stochastic switching and the Dirichlet problem for the Bellman equation. *Transactions of the American Mathematical Society*, 253:365–389, 1979.

[28] A. Fahim, N. Touzi, X. Warin, et al. A probabilistic numerical method for fully nonlinear parabolic PDEs. *The Annals of Applied Probability*, 21(4):1322–1364, 2011.

[29] M. Falcone and R. Ferretti. *Semi-Lagrangian approximation schemes for linear and Hamilton—Jacobi equations.* SIAM, 2013.

[30] X. Fang, S. Peng, Q.-M. Shao, Y. Song, et al. Limit theorems with rate of convergence under sublinear expectations. *Bernoulli*, 25(4A):2564–2596, 2019.

[31] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25. Springer Science & Business Media, 2006.

[32] V. Henderson and G. Liang. Pseudo linear pricing rule for utility indifference valuation. *Finance and Stochastics*, 18(3):593–615, 2014.

[33] V. Henderson and G. Liang. A multidimensional exponential utility indifference pricing model with applications to counterparty risk. *SIAM Journal on Control and Optimization*, 54(2):690–717, 2016.

[34] M. Hu and X. Li. Independence under the G-expectation framework. *Journal of Theoretical Probability*, 27(3):1011–1020, 2014.

[35] Y. Hu, P. Imkeller, M. Müller, et al. Utility maximization in incomplete markets. *The Annals of Applied Probability*, 15(3):1691–1712, 2005.

[36] H. Ishii. Uniqueness of unbounded viscosity solution of Hamilton-Jacobi equations. *Indiana University Mathematics Journal*, 33(5):721–748, 1984.

[37] H. Ishii. Perron's method for Hamilton-Jacobi equations. *Duke Mathematical Journal*, 55(2):369–384, 1987.

[38] H. Ishii. On uniqueness and existence of viscosity solutions of fully nonlinear second-order elliptic PDE's. *Communications on pure and applied mathematics*, 42(1):15–45, 1989.

[39] E. R. Jakobsen. On the rate of convergence of approximation schemes for Bellman equations associated with optimal stopping time problems. *Mathematical Models and Methods in Applied Sciences*, 13(05):613–644, 2003.

[40] M. Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *Annals of probability*, 28(2):558–602, 2000.

[41] N. Krylov. Nonlinear parabolic and elliptic equations of the second order. *Reidel Publishing Company Dordrecht*, 1987.

[42] N. Krylov. On the rate of convergence of finite-difference approximations for elliptic isaacs equations in smooth domains. *Communications in Partial Differential Equations*, 40(8):1393–1407, 2015.

[43] N. V. Krylov. On the rate of convergence of finite-difference approximations for Bellman's equations. *St. Petersburg Mathematical Journal*, 9(3):639–650, 1997.

[44] N. V. Krylov. On the rate of convergence of finite-difference approximations for Bellman's equations with variable coefficients. *Probability Theory & Related Fields*, 117(1), 2000.

[45] N. V. Krylov. On shige peng's central limit theorem. *Stochastic Processes and their Applications*, 130(3):1426–1434, 2020.

[46] P.-L. Lions. Optimal control of diffusion processes and hamilton–jacobi–bellman equations part 2: viscosity solutions and uniqueness. *Communications in partial differential equations*, 8(11):1229–1276, 1983.

[47] P.-L. Lions and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis*, 16(6):964–979, 1979.

[48] T. J. Lyons. Uncertain volatility and the risk-free synthesis of derivatives. *Applied mathematical finance*, 2(2):117–133, 1995.

[49] G. I. Marchuk. Some application of splitting-up methods to the solution of mathematical physics problems. *Aplikace matematiky*, 13(2):103–132, 1968.

[50] S. Nadtochiy and T. Zariphopoulou. An approximation scheme for solution to the optimal investment problem in incomplete markets. *SIAM Journal on Financial Mathematics*, 4(1):494–538, 2013.

[51] S. Peng. Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation. *Stochastic Processes and their Applications*, 118(12):2223–2253, 2008.

[52] S. Peng. A new central limit theorem under sublinear expectations. *arXiv preprint arXiv:0803.2656*, 2008.

[53] S. Peng. Backward stochastic differential equation, nonlinear expectation and their applications. In *Proceedings of the International Congress of Mathematicians 2010 (ICM 2010) (In 4 Volumes) Vol. I: Plenary Lectures and Ceremonies Vols. II–IV: Invited Lectures*, pages 393–432. World Scientific, 2010.

[54] S. Peng. Nonlinear expectations and stochastic calculus under uncertainty. *arXiv preprint arXiv:1002.4546*, 24, 2010.

[55] H. Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61. Springer Science & Business Media, 2009.

[56] A. Picarelli and C. Reisinger. Duality-based a posteriori error estimates for some approximation schemes for optimal investment problems. *Computers & Mathematics with Applications*, 79(7):2099–2118, 2020.

[57] A. Picarelli and C. Reisinger. Probabilistic error analysis for some approximation schemes to optimal control problems. *Systems & Control Letters*, 137:104619, 2020.

[58] A. Picarelli, C. Reisinger, and J. R. Arto. Some regularity and convergence results for parabolic Hamilton-Jacobi-Bellman equations in bounded domains. *Journal of Differential Equations*, 268(12):7843–7876, 2020.

[59] R. Rouge and N. El Karoui. Pricing via utility maximization and entropy. *Mathematical Finance*, 10(2):259–276, 2000.

[60] Y. Song. Normal approximation by stein's method under sublinear expectations. *Stochastic Processes and their Applications*, 130(5):2838–2850, 2020.

[61] D. W. Stroock. *Probability theory: an analytic view.* Cambridge university press, 2010.

[62] X. Tan et al. A splitting method for fully nonlinear degenerate parabolic PDEs. *Electronic Journal of Probability*, 18, 2013.

[63] A. Tourin. Splitting methods for Hamilton-Jacobi equations. *Numerical Methods for Partial Differential Equations: An International Journal*, 22(2):381–396, 2006.

[64] N. Touzi. *Optimal stochastic control, stochastic target problems, and backward SDE*, volume 29. Springer Science & Business Media, 2012.

[65] J. Vorbrink. Financial markets with volatility uncertainty. *Journal of Mathematical Economics*, 53:64–78, 2014.

[66] L. Wang. On the regularity theory of fully nonlinear parabolic equations: II. *Communications on pure and applied mathematics*, 45(2):141–178, 1992.

[67] T. Zariphopoulou. A solution approach to valuation with unhedgeable risks. *Finance and stochastics*, 5(1):61–82, 2001.